2008

MECHANISTIC INVESTIGATIONS OF THE TRANS EXCISION-SPLICING AND TRANS INSERTION-SPLICING REACTION

Perry Patrick Dotson II
University of Kentucky, ppdots2@uky.edu

ABSTRACT OF DISSERTATION

Perry Patrick Dotson II

The Graduate School
University of Kentucky
2008

MECHANISTIC INVESTIGATIONS OF THE TRANS EXCISION-SPLICING AND TRANS INSERTION-SPLICING REACTION

ABSTRACT OF DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the College of Arts and Sciences at the
University of Kentucky

By

Perry Patrick Dotson II

Lexington, Kentucky

Director: Dr. Stephen M. Testa, Associate Professor of Chemistry

Lexington, Kentucky

2008

ABSTRACT OF DISSERTATION

MECHANISTIC INVESTIGATIONS OF THE TRANS EXCISION-SPLICING AND TRANS INSERTION-SPLICING REACTION

Group I intron-derived ribozymes are catalytic RNAs that have been engineered to catalyze a variety of different reactions, in addition to the native self-splicing reaction. One such ribozyme, derived from a group I intron of *Pneumocystis carinii*, can modify RNA transcripts through either the excision or insertion of RNA sequences. These reactions are mediated through the trans excision-splicing (TES) or trans insertion-splicing (TIS) reaction pathways. To increase our current understanding of these reactions, as well as their general applicability, a mechanistic and kinetic framework for the TES reaction was established. Furthermore, additional ribozymes were investigated for their ability to catalyze the TES reaction. Lastly, the development of the TIS reaction into a viable strategy for the manipulation of RNA transcripts was investigated.

The TES reaction proceeds through two reaction steps: substrate cleavage followed by exon ligation. Mechanistic studies revealed that substrate cleavage is catalyzed by the 3' terminal guanosine of the *Pneumocystis* ribozyme. Moreover, kinetic studies suggest that a conformational change exists between the individual reaction steps. Intron-derived ribozymes from *Tetrahymena thermophila* and *Candida albicans* were also investigated for their propensity to catalyze the TES reaction. The results showed that each ribozyme could catalyze the TES reaction; however, *Pneumocystis carinii* is the most effective using the model constructs.

Investigations of the TIS reaction focused on developing a new strategy for the insertion of modified oligonucleotides into an RNA substrate. These studies used oligonucleotides with modifications to the sugar, base, and backbone positions. Each of the modified oligonucleotides was shown to be an effective TIS substrate. These results demonstrate that TIS is a viable strategy for the incorporation of modified oligonucleotides, of varying composition, into an intended RNA target.

The results from these studies show that group I introns are highly adaptable for catalyzing non-native reactions, including the TES and TIS reactions. Furthermore, group I introns are capable of catalyzing these unique reactions through distinct reaction pathways. Overall, these results demonstrate that group I introns are multi-faceted catalysts.

KEYWORDS: RNA, Group I intron, ribozyme, trans excision-splicing reaction, trans insertion-splicing reaction

Perry Patrick Dotson II

May 01, 2008

MECHANISTIC INVESTIGATIONS OF THE TRANS EXCISION-SPLICING AND TRANS INSERTION-SPLICING REACTION

By

Perry Patrick Dotson II

Stephen M. Testa, Ph.D.
Director of Dissertation

Robert B. Grossman, Ph.D.
Director of Graduate Studies

May 01, 2007

RULES FOR THE USE OF DISSERTATIONS

Unpublished dissertations submitted for the Doctor's degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with the permission of the author, and with the usual scholarly acknowledgments.

Extensive copying or publication of the dissertation in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky

A library that borrows this dissertation for use by its patrons is expected to secure the signature of each user.

Name Date

DISSERTATION

Perry Patrick Dotson II

The Graduate School
University of Kentucky
2008

MECHANISTIC INVESTIGATIONS OF THE TRANS EXCISION-SPLICING AND TRANS INSERTION-SPLICING REACTION

DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the College of Arts and Sciences at the
University of Kentucky

By

Perry Patrick Dotson II

Lexington, Kentucky

Director: Dr. Stephen M. Testa, Associate Professor of Chemistry

Lexington, Kentucky

2008

ACKNOWLEDGEMENTS

There are many people that I would like to thank who have given their time and support during my graduate career. I would like to thank my research advisor, Dr. Stephen Testa. Dr. Testa has challenged me throughout my graduate career, and has made me a more efficient scientist. I would also like to thank the members of my doctoral committee; Dr. DeMoll, Dr. Lovell, Dr. Peterson, and my outside examiner, Dr. Hunt, for their time and help throughout my graduate career. I would also like to thank the members of the Testa lab, including Dr. Rashada Alexander, Dr. Ashley Johnson, Dr. Dana Baum, Dr. Joy Sinha, Nick Tzouanakis, and Dustin Lafferty. I would like to thank Dr. Rashada Alexander and Dr. Dana Baum for showing me the techniques that I would need to get started in the lab. I would especially like to thank Dr. Joy Sinha and Nick Tzouanakis for always making things in the lab interesting.

I would like to take this opportunity to thank the people who have supported me the most. I would like to thank my parents, Perry Patrick and Lorena Lynn Dotson. I appreciate everything that they have done for me throughout my entire life. They have always been supportive of my decisions, and I am eternally grateful to have them as my parents. Without them, I would not be where I am today.

I would also like to thank my grandparents, including Michael Louis and Margaret Curry Dotson, and Walter Glenn and Lorena May Davis for their support. Although both of my grandfathers and one of my grandmothers are gone, they are not forgotten. In addition, I would like to thank all of my family members, including the Dotson's, Davis', Donohue's, Cipoletti's, and Potter's. Lastly, I would like to thank Larry and Erin White for being good friends.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE-SUMMARY

Ribonucleic acids, or RNA, are vital biomolecules in that they play a pivotal role in the pathway by which the genetic information stored within a cell is expressed. Prior to the 1980's, the role of RNA was thought to be an intermediate in the pathway of expression of proteins from deoxyribonucleic acids (DNA). The discovery of catalytic RNA resulted in the reclassification of RNA (1, 2). These discoveries ultimately led to the identification of many catalytic RNAs which are found in a variety of differing organisms. These include plants, bacteria, viruses, lower eukaryotes, and a few mammals (3-7). In the last twenty-five years, the manipulation of these catalytic RNAs, termed ribozymes, has shown that they can perform a variety of enzymatic reactions previously thought to be catalyzed only by protein enzymes (3, 6).

One previously characterized ribozyme-mediated reaction, catalyzed by a group I intron-derived ribozyme from the fungal pathogen *Pneumocystis carinii*, results in the site-specific excision of predefined regions from within RNA transcripts (8). The reaction, termed trans excision-splicing (TES), has previously been shown to be a viable strategy for the manipulation of RNA sequences *in vitro*, as well as *in vivo* (8, 9). The manipulation of RNA sequences, particularly *in vivo*, could potentially lead to the development of TES into a therapeutic strategy for the correction of mutated RNA transcripts (9). The TES reaction also provides a unique model system for the study of RNA catalysis.

Another recently characterized ribozyme-mediated reaction, termed trans insertion-splicing (TIS), has been developed for the site-specific insertion of oligonucleotides into RNA transcripts (10). Similar to the TES reaction, the TIS reaction is catalyzed by a group I intron-derived ribozyme from *Pneumocystis carinii* (10). The TIS reaction has potential as a strategy for the manipulation of RNA sequences, specifically the insertion of modified oligonucleotides into RNA transcripts.

Previous investigations of the TES reaction have led to questions relating to the overall reaction mechanism, specifically regarding the first reaction step. The research presented herein investigates the fine details of the TES reaction through mechanistic as well as kinetic characterization of the first reaction step of TES. Furthermore, ribozymes

other than that from *Pneumocystis carinii* were investigated for their ability to catalyze the TES reaction. With respect to the TIS reaction, the potential of TIS being developed into a viable strategy for the insertion of modified oligonucleotides was investigated to ascertain its feasibility.

Investigation of the Reaction Mechanism for the Trans Excision-Splicing Reaction

Previous investigations of the TES reaction suggest the first reaction step (substrate cleavage) was catalyzed through ribozyme-mediated 5' hydrolysis (8). An alternative explanation is that an endogenous nucleotide, specifically one at the 3' terminal position of the ribozyme, acts as an intramolecular nucleophile. In this case, substrate cleavage is catalyzed through ribozyme-mediated intramolecular transesterification. To distinguish between the two possible reaction mechanisms, the unique reaction intermediates that would be produced in each mechanism were analyzed.

The results demonstrate that the substrate cleavage reaction of TES is catalyzed through ribozyme-mediated intramolecular transesterification. Specifically, the 3' terminal guanosine of the ribozyme acts as an intramolecular nucleophile. This conclusion is based on several lines of evidence, including the isolation of TES ribozyme intermediates from previously characterized *in vitro* and *in vivo* model systems. Sequencing of TES ribozyme intermediates demonstrate a portion of the RNA substrate becomes covalently attached to the 3'-end of the ribozyme. Surprisingly, TES ribozyme intermediates isolated *in vivo* demonstrate that the ribozyme undergoes a self-truncation process, thereby acquiring a reactive 3' terminal guanosine. These results have greatly enhanced our current understanding of the overall mechanism for the TES reaction. For future TES studies, the mechanistic insights gained will potentially lead to an increase in the likelihood of successful TES reactions conducted *in vitro* and *in vivo*.

Kinetic Characterization of the First Step of the TES Reaction

Previous investigations of ribozyme-mediated reactions have established kinetic frameworks for gaining insight into the fundamental aspects of RNA catalysis. These previous studies have been informative with regard to the chemical and mechanistic

aspects of ribozyme-mediated reactions. Previously, a possible mechanistic pathway for the first and second reaction steps of the TES reaction has been proposed based on our current understanding of other ribozyme-mediated reactions (8, 11-17). Prior TES reports, however, have not focused on the individual reaction steps in detail, and rate constants for individual steps have not been determined. Thus, a minimal kinetic scheme, in conjunction with another member of the Testa lab (Joy Sinha), was established for the first step of the TES reaction. These studies were undertaken to increase our fundamental understanding of the mechanism for TES, as well as for RNA catalysis in general.

The results obtained from the kinetic scheme for the first reaction step (substrate cleavage) showed that the rate constants for both substrate cleavage and substrate dissociation are comparable and that the true rate for the chemical step is being masked by a conformational change. Additionally, the relatively slow rate of product dissociation suggests that it is rate-limiting with respect to the possibility of multiple-turnover reactions. The insights obtained have greatly enhanced our current understanding of the fine details with respect to the first step of the TES reaction.

The Trans Excision-Splicing Reaction Can be Catalyzed by Group I Intron-Derived Ribozymes from Tetrahymena thermophila and Candida albicans

Previous investigations of the TES reaction have focused on the group I intron-derived ribozyme from *Pneumocystis carinii* (8, 9, 18, 19). Structurally, previously characterized intron-derived ribozymes are similar to *Pneumocystis carinii* in that all the fundamental structural elements required to perform the TES reaction are present. This suggests that other group I intron-derived ribozymes could potentially catalyze the TES reaction. If true, it would be informative to know the relative effectiveness of these ribozymes. To address this possibility, the group I intron-derived ribozymes from *Tetrahymena thermophila* and *Candida albicans* were modified to investigate their ability to catalyze the TES reaction.

The results demonstrate that the intron-derived ribozymes from *Tetrahymena thermophila* and *Candida albicans* can catalyze the excision of a single nucleotide from within respective RNA substrates. In addition, both ribozyme constructs catalyze the substrate cleavage reaction through ribozyme-mediated intramolecular transesterification.

Both ribozyme constructs, however, show decreased efficiencies for catalyzing the individual steps of the TES reaction compared to the *Pneumocystis carinii* ribozyme. This was determined to be due to the individual ribozyme's ability to effectively utilize their 3' terminal guanosine as an intramolecular nucleophile. Nevertheless, our results show that the TES reaction can be catalyzed by at least three different group I intron-derived ribozymes. Apparently, being able to catalyze the TES reaction is a common property of group I intron-derived ribozymes.

Insertion of Modified Oligonucleotides using the Trans Insertion-Splicing Reaction

The site-specific modification of RNA transcripts has lead to more in-depth studies of the structural and functional aspects of RNA. Current methods for synthesizing full-length modified RNA transcripts, however, have been problematic, in some cases suffering from a lack of product specificity and low product yields. The TIS reaction was recently developed to site-specifically insert oligonucleotides into a RNA substrate (10). The next step is to determine whether the TIS reaction could be developed into a new methodology for the insertion of modified oligonucleotides into RNA transcripts. To test the feasibility of this strategy, a test system was developed that uses modified oligonucleotides as TIS substrates, including substrates with single and double modifications. For this proof-of-principle study, five modifications were used, including those to the ribose sugar (deoxy or methoxy substitution), base (2-aminopurine or 4-thiouridine), or phosphodiester backbone (thiol substitution).

The results demonstrate that the TIS reaction is a viable strategy for the insertion of modified oligonucleotides into RNA targets. For the five modified oligonucleotides tested, each was inserted sequence-specifically into their intended RNA targets, and yields ranged from 60-70%. Furthermore, the TIS strategy was shown to be effective at inserting oligonucleotides with double modifications. These results are encouraging in that the TIS strategy is potentially a new method for the synthesis of full-length modified RNA transcripts.

CHAPTER TWO-BACKGROUND

The Central Dogma of Molecular Biology.

In 1958 (20), Francis Crick proposed 'the central dogma of molecular biology' to describe the flow of genetic information within a cell (Figure 2.1). In all known cell types, the genetic information is found in the form of deoxyribonucleic acids, or DNA. For the information contained within cellular DNA to be converted into something useful, the DNA is first used as a template for the production of ribonucleic acids, or RNA. This process, known as transcription, results in a RNA transcript that is identical to the sequence of the DNA except for the substitution of thymine for uracil (see following sections). The resulting RNA transcript is further modified post-transcriptionally to yield the mature RNA species.

In the case of messenger RNA, the resulting RNA transcript is subsequently used as a template for the production of a protein product through the process of translation. Messenger RNA transcripts are divided into groups of three consecutive RNA residues, termed codons, that code for the twenty naturally occurring amino acids. During translation, each amino acid is subsequently added to the elongating polypeptide chain, through the formation of peptide bonds. The resulting protein product is further modified post-translationally, in some instances, to yield the final protein product. In later studies, RNA was also found to be used as a template for the synthesis of a complementary DNA molecule. These later discoveries, including reverse transcription, resulted in the modification of the overall scheme of the central dogma (21).

Nucleic Acids

Nucleic acids are naturally occurring biopolymers consisting of the monomeric subunits termed nucleotides. The two classes of nucleic acids, deoxyribonucleic and ribonucleic acids differ structurally, which contribute to their unique chemical and functional properties.

Deoxyribonucleic Acid

Deoxyribonucleic acid, or DNA, consists of a 2'-deoxyribose sugar with three distinct components attached (Figure 2.2A). These include a 3' hydroxyl group, a heterocyclic nucleobase, and a 5' phosphate group (22). DNA can be further differentiated into four distinct nucleotides by the identity of the heterocyclic nucleobase. Purines, which include adenine (A) and guanine (G), and pyrimidines, which include thymine (T) and cytosine (C) (Figure 2.3). Note that purines are monocyclic compounds while pyrimidines are bicyclic compounds.

Individual nucleotides are attached through phosphodiester bond linkages. This occurs through the covalent attachment of a phosphate group between the 5' hydroxyl of one nucleotide to the 3' hydroxyl of an adjacent nucleotide. Note, the resulting phosphodiester backbone of DNA has directionality in the 5' to 3' direction (Figure 2.4). Additionally, DNA is typically encountered in a double-stranded form, with each strand being oriented in an anti-parallel fashion. The two strands are complementary to one another with each type of base on one strand forming a bond with only one type of base on the other strand (22). The bonding, termed base-pairing, between the individual DNA nucleobases is through hydrogen bonding (Figure 2.5). Base-pairing typically occurs between purines and pyrimidines resulting in the normal Watson-Crick base-pairs.

Ribonucleic Acid

Ribonucleic acid, or RNA, consists of a ribose sugar ring with three separate components attached to the sugar ring (Figure 2.2B). These include a 3' hydroxyl group, a heterocyclic nucleobase, and a 5' phosphate group (22) Similar to DNA, RNA can be differentiated into four distinct nucleotides by the identity of the heterocyclic nucleobase. These include purines and pyrimidines (22). RNA differs from DNA, however, in that the 2' position of the ribose sugar is a hydroxyl group (Figure 2.2B). Additionally, the nucleobase identity is changed from uracil (U) instead of thymine (T), which differs in that a methyl group is replaced by hydrogen within the nitrogenous ring of thymine (Figure 2.3). Lastly, RNA is typically encountered in a single-stranded form having directionality in the 5' to 3' direction.

That RNA is typically encountered in a single-stranded form renders RNA capable of forming complex three-dimensional structures. RNA folding results in base-pairing between complementary regions within the RNA itself. These base-paired regions make-up the secondary, and in some cases tertiary, structures encountered for RNA. Examples of a few of the secondary structural elements that can be encountered for RNA species include duplexes, hairpins, bulges, loops, and junctions (Figure 2.6). These secondary structures, which contribute to the final tertiary structure, result in the complex shapes that RNA can acquire to perform various functions *in vivo*, including catalysis.

Catalytic RNA

Until the early 1980's, protein enzymes were thought to be the only biological species capable of performing enzymatic reactions. In 1982, an RNA species was demonstrated to be capable of catalyzing the cleavage-ligation of phosphodiester bonds in the absence of proteins (1). In this case, the intervening sequence within the large ribosomal RNA (rRNA) subunit of *Tetrahymena thermophila* was found to act as a catalyst. The autoexcision reaction, catalyzed by these introns, was later termed the self-splicing reaction. Simultaneously, the active site of ribonuclease P, isolated from *Escherichia coli* and *Bacillus subtilis*, was determined to be composed of RNA (2). Since the discovery of catalytic RNAs, also called ribozymes, several catalytic RNA species have been discovered. These include the hammerhead (23), hairpin (24), hepatitis delta virus (25), Varkud satellite (26), group I and group II introns (1, 27-29), ribonuclease P (2), and most recently the *glm*S ribozyme (30).

RNA Splicing

For some prokaryotic and eukaryotic cells, exon sequences are interrupted by an intervening sequence (IVS), or intron. Prior to the formation of a mature RNA species, the intron sequence is excised from within their flanking exon sequences. Currently, four major classes of introns have been described based upon their mechanism of excision (31, 32). These include group I and group II introns, tRNA introns, and the spliceosomal introns found within nuclear pre-mRNA (31, 32). For many group I and group II introns, the introns themselves catalyze their own excision through separate self-splicing reaction

pathways (31, 32). Note, for other group I and group II introns, proteins are required to facilitate their splicing (1, 33). As for tRNA introns, the introns are excised through enzymatic cleavage and ligation of the precursor RNA (31, 32). In the case of spliceosomal introns, the intron is excised through the complex machinery of the spliceosome (34).

Group I Introns

Group I introns have been found in the precursor mRNA, tRNA, and rRNA transcripts in a variety of different cell types, including prokaryotes, yeast, fungi, and viruses (3, 35). To date, more than 2000 group I introns are known (35). They range in size from a few hundred nucleotides to approximately 3000 nucleotides (35). In general, group I introns are composed of a common core of secondary structures, consisting of 10 to 17 base-paired regions termed P1-P17 (38-39). Upon folding, these base-paired regions constitute the active form of the intron (Figure 2.7).

Specific regions within group I introns are employed for binding and alignment of substrates (5' and 3' exon sequences) and cofactors used during the self-splicing reaction. Molecular recognition elements (REs) are employed to bind the substrates for the subsequent self-splicing reaction. The 5' exon region is bound through recognition element 1 (RE1) to form the P1 helix (40-42) (Figure 2.8). The 3' exon region is bound through recognition element 3 (RE3) to form helix P10 (40). The intron binds itself intramolecularly through recognition element 2 (RE2) to form the P9.0 helix (43, 44). Note the nucleotides that encompass RE1 and RE3, which are part of the P1 and P10 helices, are termed the internal guide sequence (IGS), in some instances. A specific region, within the P7 region of the intron, termed the guanosine-binding site (GBS) binds the two consecutive guanosine residues used in the self-splicing reaction (see section titled The Self-Splicing Reaction) (45-47).

Specific reaction sites within the exon and intron sequences define the positions of splicing, termed splice sites. For the self-splicing reaction, two splice sites (the 5' and 3' splice sites) determine the location where the two consecutive transesterification reactions of the self-splicing reaction occur. The 5' splice site is defined as the phosphodiester backbone that is 3' to the last nucleotide of the 5' exon (Figure 2.8). A

uridine residue is typically (cytosine being the only exception) the last nucleotide of the 5' exon and forms a wobble base-pair (G•U) to a conserved guanosine within RE1 of the intron (48-53). The 3' splice site is defined as the phosphodiester backbone 3' to a conserved guanosine residue located at the 3' terminal position of the intron, termed ωG.

Group I introns are also metalloenzymes in that they require metal ions to facilitate the proper folding of the intron into its active form (54-56). In addition, metal ions have been implicated in aiding in the individual steps of the self-splicing reaction. Divalent metal ions, especially magnesium (Mg^{2+}), are the preferred ion. Monovalent metal ions, including sodium (Na^{+}) or potassium (K^{+}), are acceptable, albeit at higher concentrations (54-56).

The Self-Splicing Reaction

The self-splicing reaction consists of two consecutive transesterification reactions, ultimately resulting in the excision of the intron and ligation of the flanking 5' and 3' exons. In the first reaction step, termed substrate cleavage, the intron binds an exogenous guanosine cofactor into the GBS (Figure 2.9A) (57). This exogenous guanosine can be any of the 5' phosphorylated derivatives of guanosine triphosphate (GMP, GDP, or GTP). The 3' hydroxyl group of the guanosine performs a nucleophilic attack upon the 5' splice site, resulting in cleavage of the phosphodiester backbone between the 3'-end of the 5' exon and the 5'-end of the intron. The direct result of substrate cleavage is the covalent attachment of the exogenous guanosine to the 5'-end of the intron (Figure 2.9A) (11, 58).

Prior to the second step of the self-splicing reaction (exon ligation), a proposed conformational change occurs, resulting in the expulsion of the exogenous guanosine cofactor from the GBS (46, 47, 59). This conformational change facilitates the preferential binding of ωG of the intron. The alignment of ωG into the GBS is accomplished by formation of both the P10 and P9.0 helices (60-66). For the exon ligation reaction step, the newly formed 3' hydroxyl group of the uridine at the 3'-end of the 5' exon performs a nucleophilic attack upon the 3' splice site of the intron, resulting in the ligation of the 5' and 3' exons (Figure 2.9A).

The final excised intron product, termed the linear intervening sequence, or L-IVS, can undergo subsequent cyclization reactions that result in the intron self-processing itself into a circular form (67-70). In the cyclization reaction, the L-IVS binds a portion of the 5'-end of the intron through RE1 to form a P1 helix, resulting in the establishment of a new 5' splice site (Figure 2.9B). The L-IVS catalyzes an intramolecular transesterification reaction, resulting in the covalent attachment of the 5' and 3'-ends of the L-IVS. The intramolecular transesterification reaction is catalyzed by the 3' terminal guanosine, ωG, of the L-IVS. Overall, the reaction results in a circularized intron product, termed the circular intervening sequence, or C-IVS (Figure 2.9B).

Group I Intron-Derived Ribozymes

To further elucidate the overall mechanism of the self-splicing reaction, group I introns have been modified by removal of their native 5' and 3' exon sequences (Figure 2.9). The new catalytic RNA species, known as group I intron-derived ribozymes, were designed so that the individual steps of the self-splicing reaction could be investigated. Of these group I intron-derived ribozymes, the *Tetrahymena thermophila* ribozyme (71) has been the most extensively studied of these ribozymes. Other group I intron-derived ribozymes, however, have been constructed. These include the intron-derived ribozymes from *Anabaena* (72), *Azoarcus* (73), *Candida albicans* (74), and *Pneumocystis carinii* (75). Group I intron-derived ribozymes have been shown to be capable of reaction rate enhancements similar to enzyme catalysts and are capable of multiple-turnover reactions (71-74).

Group I intron-derived ribozymes have been shown to be capable of catalyzing additional non-native reactions, including ribonuclease (14, 15), phosphatase (13), polymerization (76, 77), recombination (78), reverse-splicing (79, 80), trans-splicing (81-90), trans excision-splicing (8), trans insertion-splicing (10), and 5' replacement reactions (91). That group I intron-derived ribozyme can catalyze additional reactions demonstrates the multi-faceted nature of these catalysts.

The Trans Excision-Splicing Reaction

The trans excision-splicing (TES) reaction, catalyzed by a group I intron-derived ribozyme from the opportunistic pathogen *Pneumocystis carinii*, was developed to sequence specifically excise a targeted sequence from within an RNA substrate (Figure 2.10) (8). The TES reaction consists of two consecutive transesterification reactions that are similar to those of the self-splicing reaction. These consist of substrate cleavage followed by exon ligation. The TES reaction has previously been shown to be capable of excising as little as one nucleotide and as many as twenty-eight nucleotides from within RNA substrates (8). Most recently, the TES reaction was shown to be an effective strategy for the manipulation of RNA transcripts *in vivo* utilizing an *Escherichia coli* model system (9).

As seen for intron-catalyzed self-splicing, TES ribozymes appear to utilize the same molecular recognition elements for orienting their substrates and intermediates during the TES reaction. The 5' exon region is bound through recognition element 1 (RE1) to form the P1 helix (Figure 2.10). The 3' exon region is bound through recognition element 3 (RE3) to form the P10 helix. For TES reactions resulting in the excision of more than one nucleotide, the bridge region is bound through recognition element 2 (RE2) to form the P9.0 helix (Figure 2.10). Though previous investigations have shown that formation of helix P10 and P9.0 are beneficial for the overall TES reaction, neither is required (18, 19).

In the first reaction step (substrate cleavage), the *Pneumocystis carinii* ribozyme mediates the cleavage of the phosphodiester backbone located at the 3'-end of the 5' exon region (Figure 2.10). Therefore, for TES, the 5' splice site is determined as the 3'-end of helix P1. Under optimal TES conditions, a G•U wobble base-pair is formed at the 5' splice site, however, other base-pair combinations have been substituted successively (19). These include Watson-Crick base-pairs, though decreased TES product yields were observed (19).

For the second reaction step (exon ligation), the nucleophilic 5' exon attacks the phosphodiester backbone 3' of a specific guanosine, termed ωG, within the substrate. The ωG, which aids in defining the 3' splice site, is defined as the border between the bridge region and the 3' exon region (Figure 2.10). Furthermore, previous investigations have

shown that only a guanosine can be defined at the ω-position of the substrate (19). The products of exon ligation consist of the final TES product (ligated 5' and 3' exon sequences) and the bridge region covalently attached to the ribozyme (Figure 2.10).

The Trans Insertion-Splicing Reaction

The trans insertion-splicing (TIS) reaction, catalyzed by a group I intron-derived ribozyme from *Pneumocystis carinii*, was recently developed to site-specifically insert a targeted sequence into an RNA substrate (Figure 2.11) (10). The TIS reaction, which utilizes two separate RNA substrates, has been proposed to proceed through three separate chemical steps. This includes two consecutive substrate cleavage steps and a single exon ligation step (Figure 2.11). As seen for the ribozyme-mediated TES reaction, TIS ribozymes appear to utilize the same molecular recognition elements for orienting their substrates and intermediates during the TIS reaction. The TIS reaction has previously been shown to be capable of catalyzing the insertion of six nucleotides into a separate RNA substrate (10).

In the first reaction step (substrate cleavage), the TIS donor substrate is bound through the first three nucleotides of Recognition Element 1 (RE1) to form a partial P1 helix, termed P1i (Figure 2.11). In addition, the TIS donor substrate is bound through Recognition Element 3 (RE3) to form a P10 helix. The 3' terminal guanosine (G336) of the *Pneumocystis carinii* ribozyme performs a nucleophilic attack upon the 5' splice site (Figure 2.11). The 5' splice site is defined as the phosphodiester backbone that is 3' to the last nucleotide of the 5' exon region. The result of the substrate cleavage reaction is the covalent attachment of the insert region of the TIS donor substrate to the 3'-end of the *Pneumocystis carinii* ribozyme. After the substrate cleavage reaction, the 5' exon region of the TIS donor substrate dissociates from the ribozyme (Figure 2.11).

Prior to the second reaction step, the ribozyme binds the 5' exon region of the TIS acceptor substrate through RE1 to form a new P1 helix (Figure 2.11). In the second reaction step (substrate cleavage), the 3' terminal guanosine of the insert region, termed ωGi, performs a nucleophilic attack upon the new 5' splice site (Figure 2.11). In this case, the 5' splice site is defined as the phosphodiester backbone that is 3' to the last nucleotide of the 5' exon region. Substrate cleavage results in the covalent attachment of

the 3’ exon region of the TIS acceptor substrate to the 3’-end of the *Pneumocystis carinii* ribozyme (Figure 2.11).

Prior to the final step of the TIS reaction, a proposed conformational change occurs, which displaces ωGi from the GBS. This results in preferential binding of G336 of the ribozyme to the GBS, which aids in defining the 3’ splice site. The alignment of G336 into the GBS is thought to occur through reformation of helix P10. In the third reaction step (exon ligation), the 3’ hydroxyl group of uridine performs a nucleophilic attack upon the phosphodiester backbone 3’ to G336. This results in final TIS product formation, including ligation of 5’ exon region to the 3’ exon region, including the incorporation of the insert (Figure 2.11).

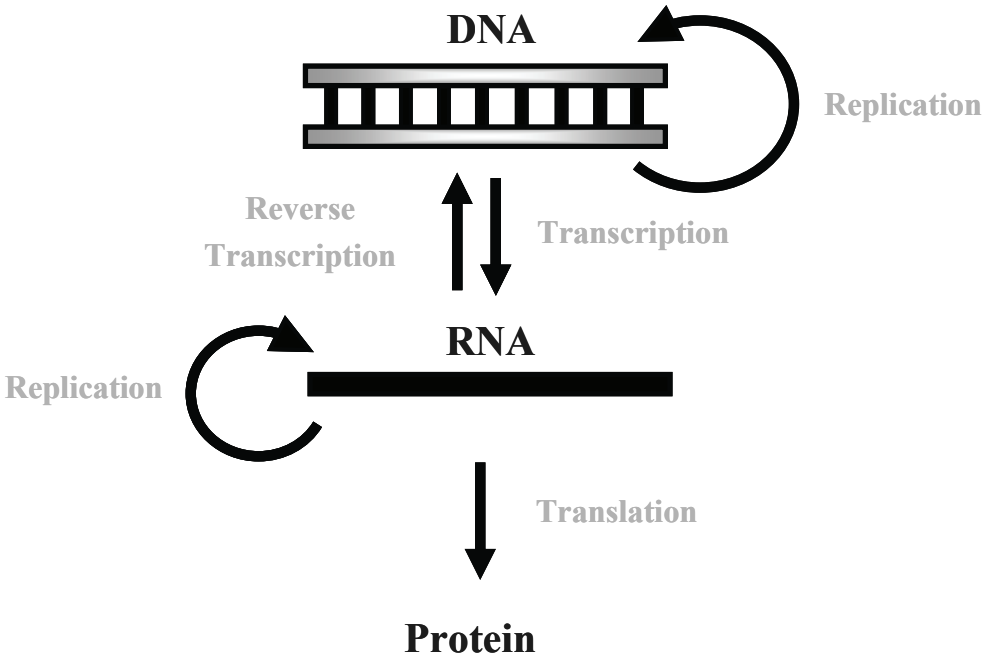


Figure 2.1. The central dogma of molecular biology. Deoxyribonucleic acids, or DNA, are the biological macromolecule that contain all the genetic information of a cell, and is typically found in a double helical form. Prior to cell replication, DNA is replicated so that an identical copy of the parental DNA is passed to each successive generation. The genetic information stored in DNA is first transcribed to ribonucleic acids, or RNA, and is then translated to protein. Additionally, RNA can be used as a template for the creation of DNA through reverse transcription.

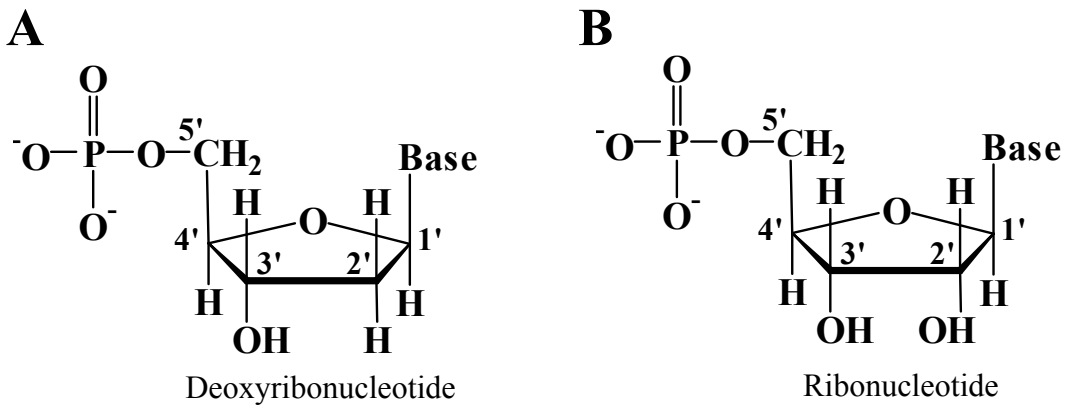


Figure 2.2. Structure of a deoxy and ribonucleotide. A. Diagram of a deoxyribonucleotide including the distinct functional groups attached to the pentose sugar ring including the nucleobase, hydroxyl, and phosphate groups. B. Diagram of a ribonucleotide including the distinct functional groups attached to the pentose sugar ring including the nucleobase, hydroxyl, and phosphate groups. Ribonucleotides differ from their deoxyribonucleotide counterparts in that a hydroxyl group is substituted at the 2' position of the sugar ring. For the case of DNA and RNA, the identity of the nucleobase distinguishes the four nucleotides.

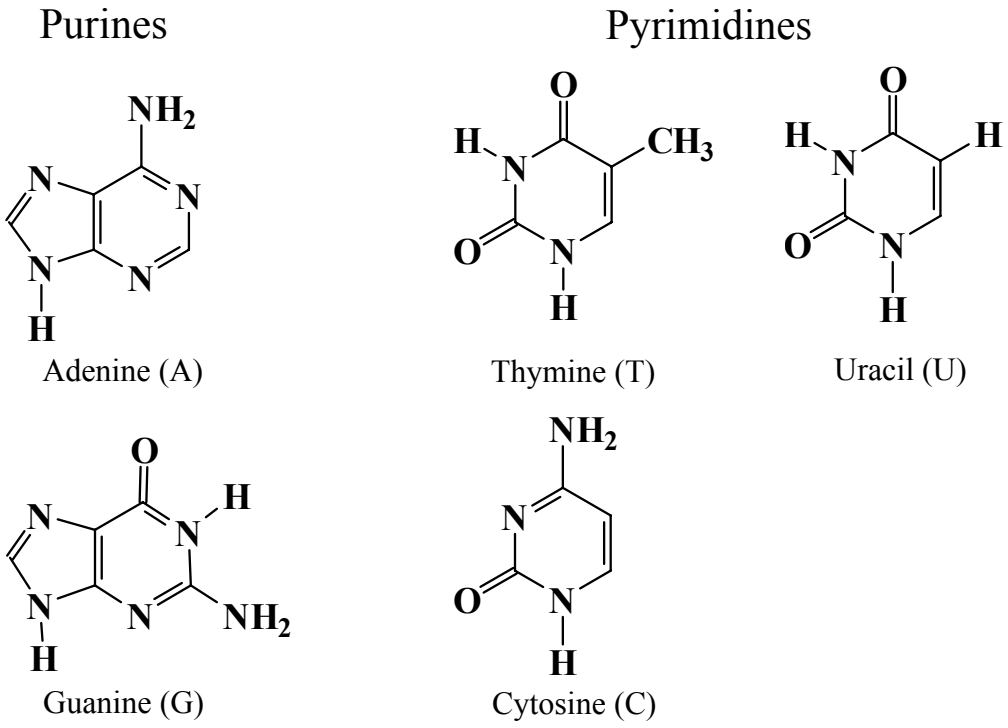


Figure 2.3. Nucleobases encountered in DNA and RNA. The nucleobases encountered for DNA and RNA are separated into bicyclic purines and monocyclic pyrimidines. Purines include adenine and guanine, and pyrimidines include thymine, uracil, and cytosine. DNA consists of the adenine, guanosine, thymine, and cytosine nucleobases. RNA consists of adenine, guanosine, uracil, and cytosine nucleobases. The individual bases are typically indicated with their single letter abbreviation; adenine (A), guanine (G), thymine (T), uracil (U), and cytosine (C). DNA and RNA sequences are typically denoted using single letter abbreviations, which represent the individual nucleotides.

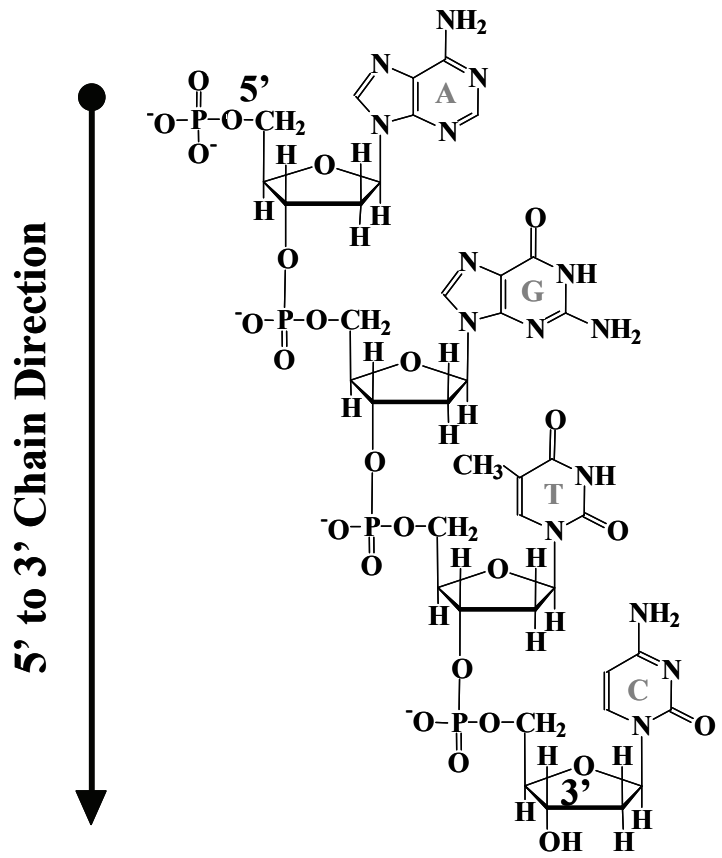


Figure 2.4. Directionality of DNA. Diagram of directionality of DNA from the 5' phosphate group to the 3' hydroxyl group depicting all four nucleotides (A = adenosine, G = guanosine, T = thymine, and C = cytosine). Each of the nucleotides is linked through a 5', 3' phosphodiester bond linkage. An arrow denotes the directionality of the DNA.

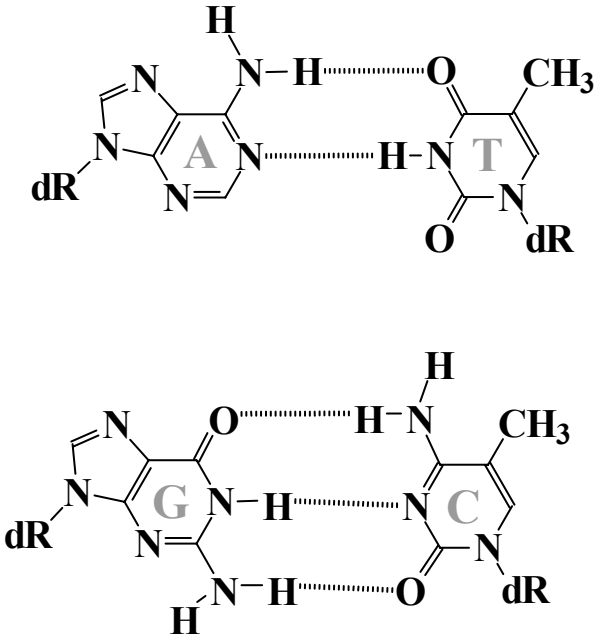


Figure 2.5. Watson-Crick base-pairing typically encountered in DNA. Diagram of Watson-Crick base-pairs found within DNA. The top pair represents the base-pairing interaction between adenine and thymine (two hydrogen bonds). For RNA, a base-pair between adenine and uracil occurs. The bottom pair represents the base pairing interaction between guanine and cytosine (three hydrogen bonds). Note, dR represents the deoxyribose sugar.

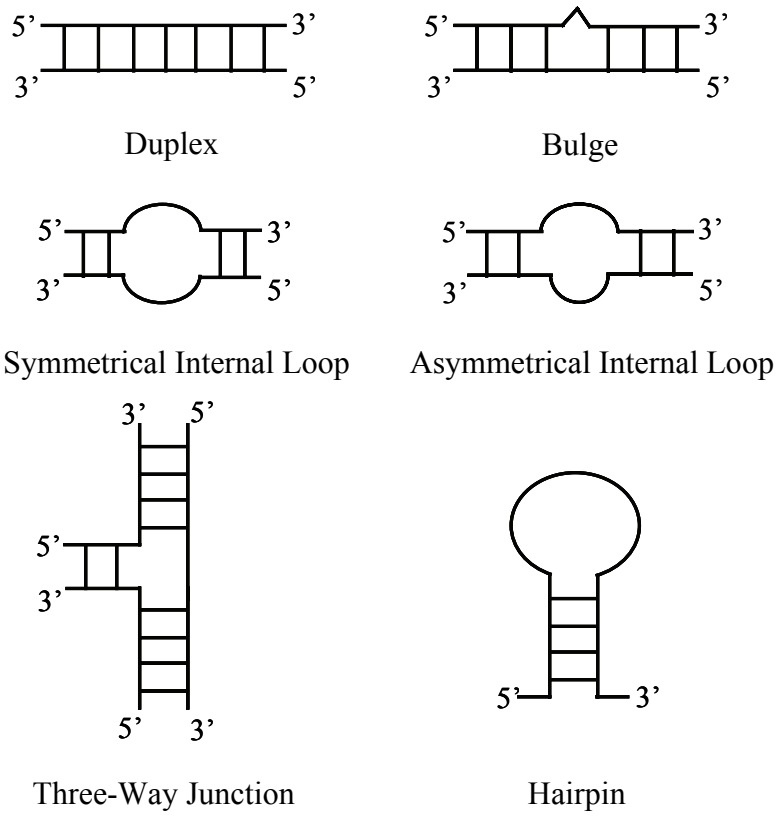


Figure 2.6. Common secondary structures encountered for RNA. Schematic representation of common secondary structures encountered for RNA including duplexes, bulges, hairpins, internal loops (both symmetrical and asymmetrical), and junctions are depicted.

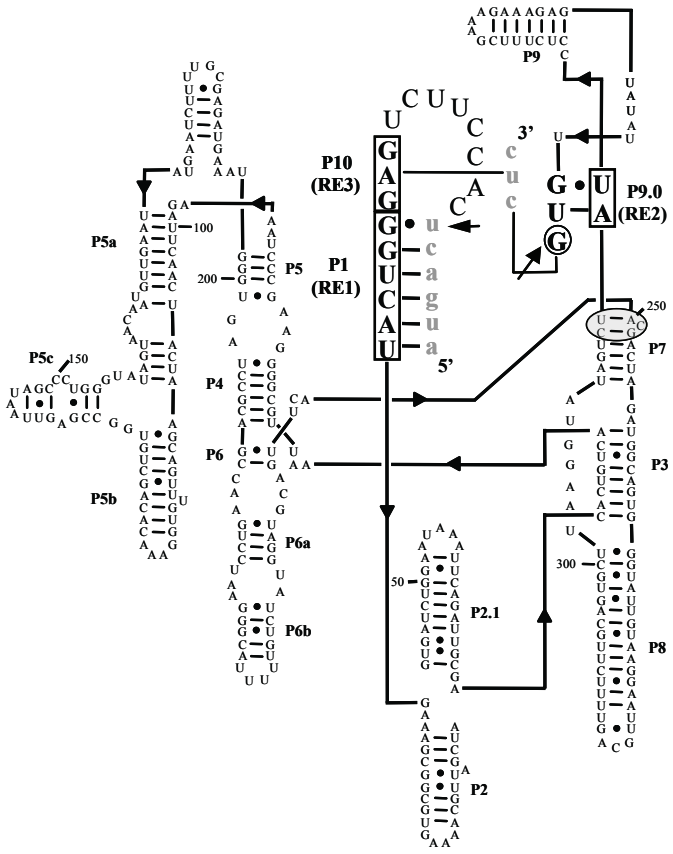


Figure 2.7. Sequence and proposed secondary structure of the *Pneumocystis carinii* group I intron. The intron sequence is denoted in black upper case lettering and the exon sequences are denoted in gray lowercase lettering. The molecular recognition elements (boxed) are shown whereby recognition element 1 (RE1) binds the 5' exon forming the P1 helix. Recognition element 3 (RE3) binds the 3' exon forming the P10 helix, and the intron binds itself intramolecularly through recognition element 2 (RE2) to form the P9.0 helix. The sites of reaction, termed splice sites, are denoted by arrows. The guanosine-binding site (GBS) is shown within the P7 region of the intron circled in gray, which binds the two consecutive guanosines utilized for the self-splicing reaction.

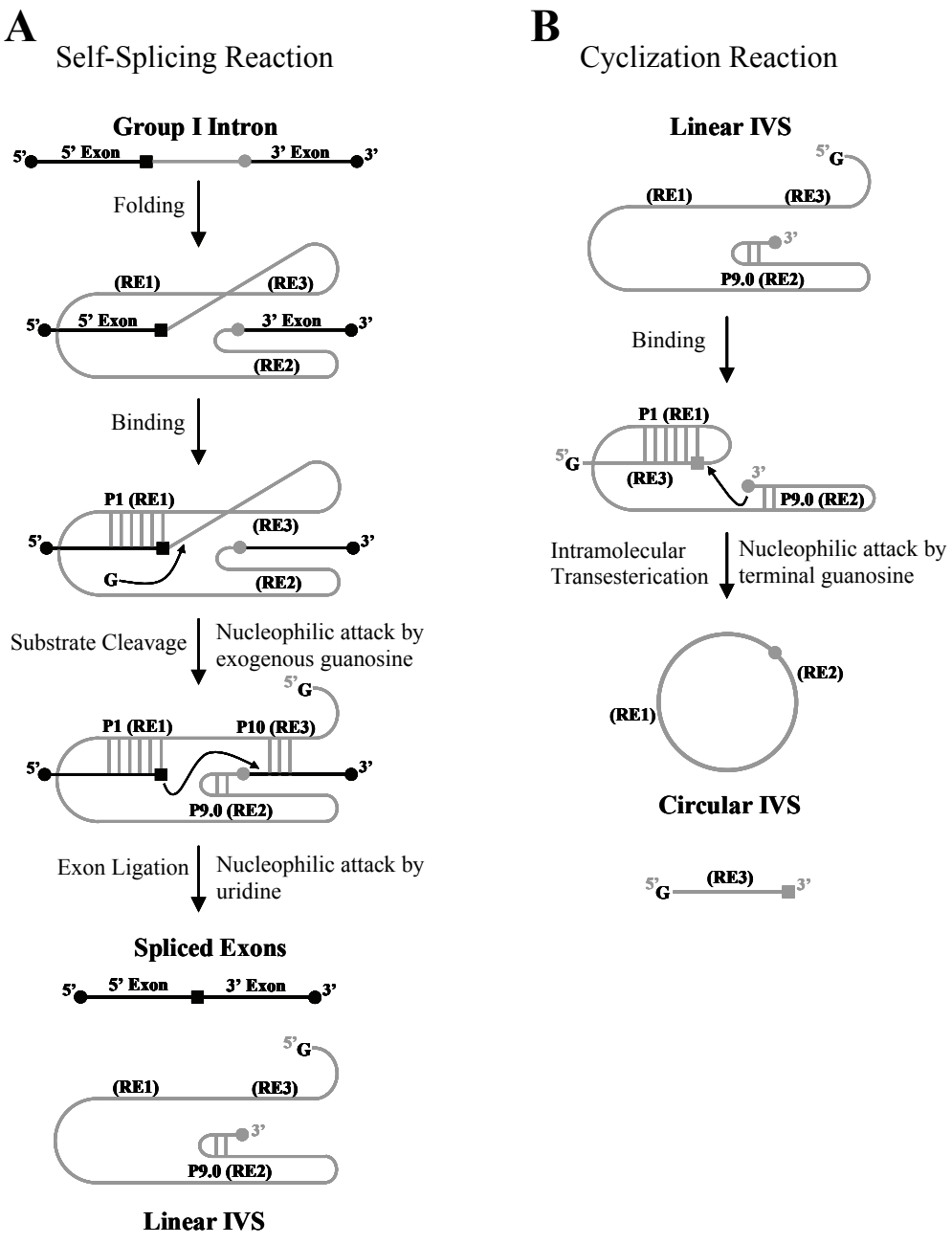


Figure 2.8. The reaction pathway for the self-splicing and cyclization reactions catalyzed by group I introns. A. The self-splicing reaction. The group I intron is denoted by a gray line and the 5' and 3' exon sequences are denoted by black lines. A black square within the 5' exon region denotes a uridine and a gray circle at the 3'-end of the intron denotes a guanosine (ωG**).** The P1 helix is formed through base-pairing of recognition element 1

(RE1) with the 5’ exon. The P10 helix is formed through base-pairing of recognition element 3 (RE3) with the 3’ exon. The intron binds itself intramolecularly through recognition element 2 (RE2) to form the P9.0 helix. The first reaction step (substrate cleavage) is catalyzed by an exogenous guanosine cofactor. Substrate cleavage results in covalent attachment of the guanosine to the 5’-end of the intron. The second step (exon ligation) proceeds through attack of the uridine upon the ωG of the intron, resulting in ligation of the flanking sequences. The resulting intron product is termed the linear intervening sequence (L-IVS). B. The cyclization (circularization) reaction catalyzed by the L-IVS. The L-IVS binds a portion of the 5’-end of the L-IVS (typically including the RE3 region) to form a new P1 helix. An intramolecular transesterification reaction, catalyzed by ωG (gray circle at 3’-end of L-IVS) occurs, resulting in the covalent attachment of the 5’-end to the 3’-end of the L-IVS. This forms the circular intervening sequence (C-IVS).

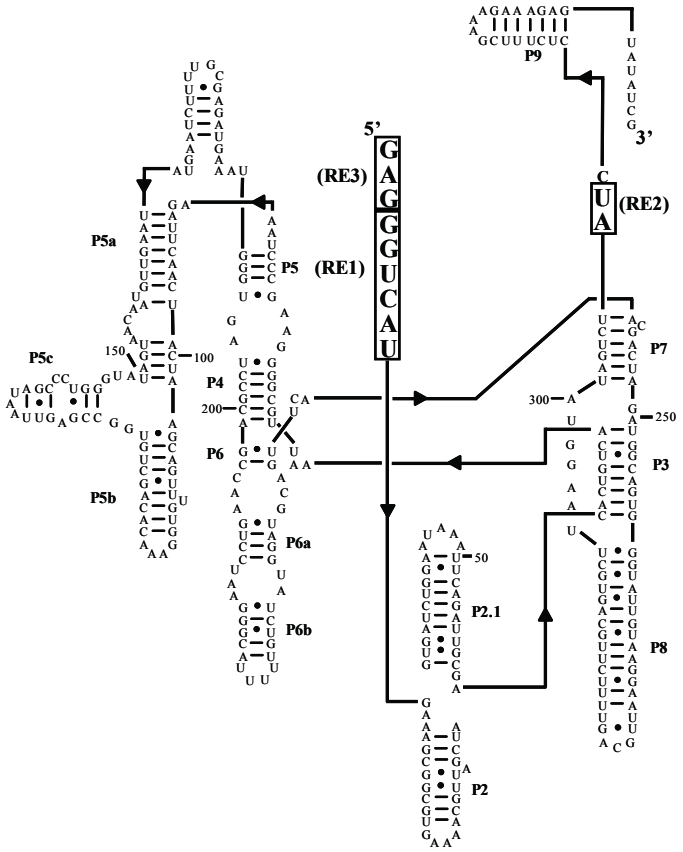


Figure 2.9. Sequence and proposed secondary structure of a *Pneumocystis carinii* group I intron-derived ribozyme. The primary sequence of the group I intron-derived ribozyme is essentially identical to the group intron depicted in Figure 2.8. The difference between the two sequences is removal of both the 5' exon and 3' exon sequences. In addition, eight nucleotides are removed from the 5'-end of the intron and four nucleotides are removed from the 3' end of the intron. The molecular recognition elements (shown in bold, boxed uppercase lettering) consist of recognition element 1 (RE1), recognition element 3 (RE3), and recognition element 2 (RE2) which will be used for formation of the P1, P10, and P9.0 helices respectively.

Trans Excision-Splicing Reaction

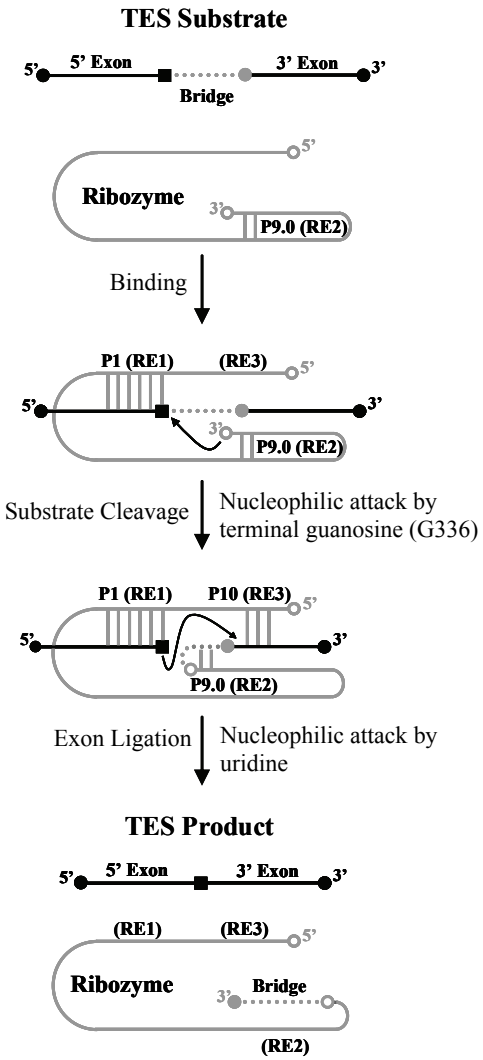


Figure 2.10. Diagram of the ribozyme-mediated trans excision-splicing reaction. The *Pneumocystis carinii* ribozyme is denoted in gray lines and the 5' and 3' exon sequences are denoted by black lines. The black square within the 5' exon region denotes a uridine and the gray circle adjacent to the 3' exon region denotes a guanosine (ωG). The bridge region, to be excised, is denoted as a dotted line. The P1 helix is formed through base-pairing of recognition element 1 (RE1) with the 5' exon. The P10 helix is formed through base-pairing of recognition element 3 (RE3) with the 3' exon. The first step (substrate cleavage) is catalyzed by G336 (open circle at 3'-end of ribozyme) resulting in the

covalent attachment of the 3'-end of the substrate to the 3'-end of the ribozyme. The second step (exon ligation) proceeds through attack of the uridine upon the ωG of the substrate resulting in ligation of the flanking sequences. Note the ribozyme product of the TES reaction is that ωG (of substrate) remains attached to the 3'-end of the ribozyme.

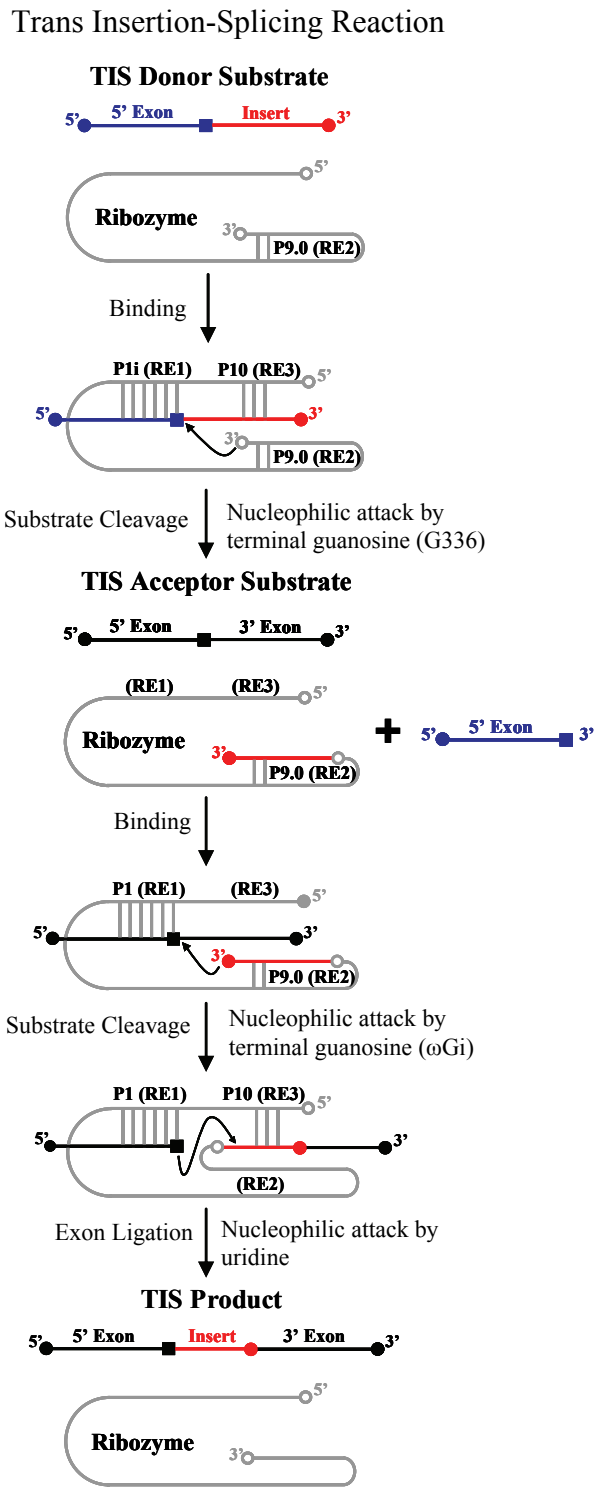


Figure 2.11. Diagram of the ribozyme-mediated trans insertion-splicing reaction. The *Pneumocystis carinii* ribozyme is denoted by gray lines and the TIS donor substrate is

separated into blue (5’ exon) and red (insert) lines. A blue square within the TIS donor substrate denotes a uridine and the red square at the 3’-end denotes a guanosine (ωGi). The TIS acceptor substrate is denoted by black lines and the black square denotes a uridine. The ribozyme recognition elements; RE1, RE2, and RE3 that form the P1, P9.0, and P10 helices are depicted. The TIS donor substrate forms the P1i and P10 helices through RE1 and RE3 of the ribozyme. In the first reaction step (substrate cleavage), the 3’ terminal guanosine (G336) of the ribozyme (open circle at the 3’-end of ribozyme) performs a nucleophilic attack upon the 5’ splice site of the TIS donor substrate. This results in the covalent attachment of the insert region of the TIS donor substrate to the 3’-end of the ribozyme. The 5’ exon region of the TIS donor substrate dissociates from the ribozyme and the TIS acceptor substrate is bound through a second P1 helix. For the second reaction step (substrate cleavage), the ωGi (aligned by P9.0 helix formation) performs a nucleophilic attack upon the 5’ splice site in the TIS acceptor substrate. This results in the covalent attachment of the 3’ exon region of the TIS acceptor substrate to the 3’-end of the ribozyme. The last reaction step (exon ligation) proceeds through the nucleophilic attack of the uridine of the 5’ exon region of the TIS acceptor substrate upon G336 of the ribozyme, resulting in the final TIS product.

CHAPTER THREE-THE GROUP I INTRON-DERIVED RIBOZYME FROM *PNEUMOCYSTIS CARINII* UTILIZES AN ENDOGENOUS GUANOSINE AS THE FIRST REACTION STEP NUCLEOPHILE FOR TRANS EXCISION-SPLICING

INTRODUCTION

The trans excision-splicing (TES) reaction, catalyzed by a group I intron-derived ribozyme from the opportunistic pathogen *Pneumocystis carinii* (*P. carinii*), has previously been shown to be capable of site-specifically excising targeted sequences from within RNA substrates (8). The TES reaction, which is similar to the self-splicing reaction, consists of two consecutive transesterification steps whereby the same molecular recognition elements employed by other group I intron-derived ribozymes are exploited for binding and alignment of subsequent TES substrates. One key difference between the reactions, however, is that the self-splicing reaction utilizes an exogenous guanosine cofactor as the substrate cleavage nucleophile (3), while TES does not (8).

In the absence of an exogenous guanosine cofactor, group I intron-derived ribozymes can exploit two separate reaction pathways to catalyze the cleavage of their intended RNA substrate. These include ribozyme-mediated 5' hydrolysis or intramolecular transesterification (Figure 3.1). It was previously postulated that the *P. carinii* ribozyme could be catalyzing the substrate cleavage reaction through ribozyme-mediated 5' hydrolysis (Figure 3.1A) (8). An alternative explanation is that an endogenous nucleotide, most likely one at the 3'-end of the ribozyme (Figure 3.1B), is the substrate cleavage nucleophile. There is precedent for both reactions (12-17). Ribozyme-mediated hydrolysis reactions have previously been shown to occur for other group I introns and their derived ribozymes, including the intron-derived ribozyme from *Tetrahymena thermophila* (*T. thermophila*) (12, 16). The *T. thermophila* group I intron has also been shown to be capable of catalyzing intramolecular cyclization (transesterification) reactions utilizing the 3' terminal guanosine of the intron as an intramolecular nucleophile (see Figure 2.1B) (13, 15, 17).

To further characterize the TES reaction, the first reaction step (substrate cleavage) was investigated by following the accumulation of TES reaction intermediates through the course of the reaction. It was found that the substrate cleavage reaction proceeds through an intramolecular transesterification reaction pathway and is catalyzed by the 3' terminal guanosine, in this case G336, of the *P. carinii* (rPC) ribozyme (Figure 3.2). The intramolecular transesterification reaction pathway is supported by several lines of evidence including the observation that the identity of the 3' terminal nucleotide is critical for optimal TES activity and the size and identity of ribozyme reaction intermediates isolated from TES reactions (both *in vitro* and *in vivo*). Additionally, *in vivo* results demonstrate that the reactive 3' terminal guanosine can be created through self-truncation of the ribozyme transcript, most likely through ribozyme-mediated 3' hydrolysis. Apparently, the ribozyme activates itself for the TES reaction in a cell. The implications of these results demonstrate that the ribozyme itself plays a more substantial role than previously thought in that the ribozyme itself is an intermediate in the TES reaction pathway.

MATERIALS AND METHODS

Oligonucleotide Synthesis and Preparation

DNA oligonucleotides were obtained from Integrated DNA Technologies (Coralville, IA) and were used without further purification. RNA oligonucleotides were obtained from Dharmacon Research, Inc. (Lafayette, CO) and were deprotected using the manufacturer's suggested protocol. Concentrations of individual oligonucleotides were calculated from UV absorption measurements using a Beckman DU 650 spectrophotometer (Beckman Coulter, Inc.; Fullerton, CA). RNA oligonucleotides were 5'-end radiolabeled by phosphorylation of the 5' terminal hydroxyl group with [γ-^{32}P] ATP (Amersham Pharmacia Biotech; Piscataway, NJ) using T4 polynucleotide kinase (New England Biolabs; Beverly, MA) as previously described (75). RNA oligonucleotides were 3'-end radiolabeled by ligation of 5'-end radiolabeled pCp (5'-*pCp) to the 3'-end of the RNA oligonucleotide using T4 RNA ligase (New England Biolabs) as previously described (10). Note that the 3'-end radiolabeling strategy will

result in substrates being a nucleotide longer than 5'-end radiolabeled substrates.

Ribozyme Synthesis and Preparation

Two ribozymes were constructed via run-off transcription of linearized plasmids. For both cases, the *P. carinii* ribozyme plasmid, PC, was linearized in a 50 μL reaction consisting of 16 μg plasmid, 5 μL of 10X REACT 2 buffer (Invitrogen; Grand Island, NY), and 50 units of either *Xba*I (to form PC-X) or *Hind*III (to form PC-H) at 37 °C for 2 h. The linearized plasmids were purified using a QIAquick PCR purification kit (Qiagen; Valencia, CA).

Two additional ribozymes were constructed via PCR using the PC plasmid as the template. The following primers were used for the individual ribozyme constructs: the forward PCR primer for both ribozyme constructs was $^{5'}$CTCTAATACGACTCACTATAGAGGG$^{3'}$, and the reverse PCR primers (variable regions are underlined) for each of the ribozyme constructs were $^{5'}$TTAGATATACTCTTTCTTTCGAAAGAGG$^{3'}$ for PC-ωA (10), and $^{5'}$CTAGATATACTCTTTCTTTCGAAAGAGG$^{3'}$ for PC-ωG. All PCR reactions utilized 45 pmol of each primer in a reaction consisting of 100 ng PC plasmid, 1 unit of Pfu DNA polymerase (Stratagene; La Jolla, CA), 1X Pfu reaction buffer (Stratagene), and 0.5 μM dNTPs. The PCR conditions were 95 °C for 30 s (initial denaturation), followed by 40 cycles of 95 °C for 30 s (denaturation), 55 °C for 45 s (annealing), and 68 °C for 120 s (elongation). A final elongation step of 68 °C for 360 s was performed after the final cycle. PCR products were separated on a 2% agarose gel and were purified using the QIAquick Gel Extraction Kit (Qiagen) using the manufacturers protocol.

Both the linearized and PCR amplified DNA templates were used for run-off transcription in a 50 μL reaction consisting of 1 μg DNA template, 50 units of T7 RNA polymerase (New England Biolabs), 40 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 5 mM DTT, 5 mM spermidine, and 1 mM rNTP mix for 2 h at 37 °C. Each of the individual ribozyme constructs were purified using a Qiagen Plasmid Midi Kit (Qiagen), as previously described (8).

TES Reactions

TES reactions were conducted at 44 °C in H10Mg buffer consisting of 50 mM HEPES (25 mM Na^+), 135 mM KCl, and 10 mM $MgCl_2$. Prior to each reaction, 200 nM ribozyme in 25 μL H10Mg buffer was preannealed at 60 °C for 5 min and then slow cooled to 44 °C. The reaction was initiated by adding 5 μL of a 10 nM solution of radiolabeled substrate (also in H10Mg solution). The final concentrations of the ribozyme and substrate in the reaction mixtures were 166 nM and 1.3 nM respectively. Time course studies were conducted by the periodic removal of 3 μL aliquots, which were then quenched by adding an equal volume of 2 X stop buffer (10 M urea, 3 mM EDTA, and 0.1 X TBE). The products and substrates were denatured at 90 °C for 1 min and then separated on a 12.5% denaturing polyacrylamide gel. Gels were dried under vacuum and the bands were visualized on a Molecular Dynamics Storm 860 Phosphoimager. The observed rate constant, k_{obs}, was obtained from a plot of the percent substrate cleavage product (substrate cleavage plus TES product) over time (8). Note the k_{obs} values obtained are reported as lower limit values, in some instances, and cannot be taken as precise values due to the limits of manual pipetting.

RT-PCR Amplification of TES Reaction Intermediates In Vitro

TES reactions were conducted using a 20-mer ($^{5'}$AUGACUACUCUCGU*G*CUCUU$^{3'}$) substrate using the conditions described above. The products were amplified via RT-PCR with the following primers: $^{5'}$GAGGGTCATGAAAGCGGCGTG$^{3'}$ (specific for the rPC ribozyme) and $^{5'}$GCACCGGTAAGAGCACGAGAG$^{3'}$ (specific for the bridge-3' exon). RT-PCR reactions were conducted using an Access RT-PCR kit (Promega; Madison, WI) and consisted of a 50 μL reaction containing approximately 2 μL TES reaction (166 nM ribozyme), 1 mM $MgSO_4$, 45 pmol of each primer, 0.2 mM dNTPs, 5 units AMV reverse transcriptase, and 5 units Tfl DNA polymerase. The reactions were subjected to first strand synthesis at 45 °C for 45 min, which was followed by 2 min at 94 °C for deactivation of the AMV reverse transcriptase. The reactions were then subjected to 40 PCR cycles consisting of 94 °C for 30 s (denaturation), 55 °C for 60 s (annealing), and 68 °C for 120 s (elongation). After the cycling was complete, a final elongation step of 68

°C for 10 min was performed. The RT-PCR products were separated on a 2% agarose gel and excised from the gel matrix using the QIAquick Gel Extraction Kit (Qiagen). The excised products were ligated into the pDrive cloning vector (Qiagen) using the manufacturer's recommended protocol and transformed into DH5α cells (Invitrogen). Colonies were picked and sequenced for identification (Davis Sequencing; Davis, CA).

In Vivo Characterization of TES Intermediates

Plasmids were used that contain both the ribozyme and GFP target transcript. In all cases, plasmids were transformed into *Escherichia coli* (*E. coli*) strain JM109 (DE3), which encodes T7 RNA polymerase, as described (9). Once transformed, T7 RNA polymerase was induced by induction with isopropylthiogalactoside (IPTG), which resulted in the transcription of both the ribozyme and GFP, as described (9). The plasmids: pQBI-Mut-GFP+p3x-RE3 (5) (active ribozyme with mutant GFP) and pQBI-Mut-GFP+p3x-RE3 (5)-ΔGBS (inactive ribozyme with mutant GFP) were used (9). After transformation and induction, 1.5 mL of the designated cell culture was used for total RNA isolation using the Ambion RiboPure^{TM}-Bacteria Kit (Ambion; Austin, TX). The TES reaction intermediate was amplified via RT-PCR using the following primers: $^{5'}$GTGGAGAGGTAGAAAGCGGCGTG$^{3'}$ (specific for *P. carinii* ribozyme) and $^{5'}$CCATGCCATGTGTAATCCCAGCAGC$^{3'}$ (specific for 3'-end of GFP transcript), which amplify products containing only both the ribozyme and the GFP transcript. RT-PCR reactions were conducted using the Access RT-PCR system (Promega) consisting of a 50 μL reaction containing approximately 1 μg total RNA, 1 mM $MgSO_4$, 45 pmol of each primer, 0.2 mM dNTPs, 5 units AMV reverse transcriptase, and 5 units Tfl DNA polymerase. Reactions were then subjected to first strand synthesis (45 °C for 45 min), which was followed by 2 min at 94 °C for deactivation of the AMV reverse transcriptase. These reactions were then subjected to 40 PCR cycles consisting of 94 °C for 30 s (denaturation), 55 °C for 60 s (annealing), and 68 °C for 120 s (elongation). After the cycling was complete a final extension step of 68 °C for 10 min was performed. The RT-PCR products were separated on a 2% agarose gel and were excised from the gel matrix using the QIAquick Gel Extraction Kit (Qiagen). The excised products were ligated into a pDrive cloning vector (Qiagen) using the manufacturer's recommended protocol and

transformed into DH5α cells (Invitrogen). Colonies were picked and sequenced (Davis Sequencing) for identification.

RESULTS

TES Reaction Intermediates In Vitro

The two possible substrate cleavage reaction mechanisms, ribozyme-mediated 5' hydrolysis or intramolecular transesterification cleavage, can be distinguished by following the reaction intermediates for TES reactions conducted *in vitro* (Figure 3.1). For TES reactions conducted with the substrate, $^{5'}$AUGACU*G*CUC$^{3'}$, the 5' hydrolysis mediated reaction will result in two reaction intermediates consisting of the 5'-exon ($^{5'}$AUGACU$^{3'}$) and bridge-3' exon ($^{5'}$*G*CUC$^{3'}$) (Figure 3.1A). In contrast, the intramolecular transesterification reaction will result in two reaction intermediates consisting of the 5'-exon ($^{5'}$AUGACU$^{3'}$) and the bridge-3'-exon being covalently attached to the 3'-end of the rPC ribozyme ($^{5'}$rPC-*G*CUC$^{3'}$) (Figure 3.1B). However, there is no size difference between the 5'-exon sequences, which precludes discriminating the two cleavage pathways using 5'-end radiolabeled substrate. Using a 3'-end radiolabeled substrate, however, could potentially distinguish if the bridge-3'-exon is separate from the ribozyme or if it is being attached to the 3'-end of the ribozyme. Therefore, we synthesized the 3'-end radiolabeled substrate ($^{5'}$AUGACU*G*CUC<u>C</u>$^{3'}$ italic lettering denotes the nucleotide to be excised and underlined lettering denotes the additional nucleotide due to 3'-end radiolabeling) and followed the formation of the intermediate and product during the normal course of the TES reaction.

TES reactions were conducted under previously optimized reaction conditions (8) using both 3' and 5'-end radiolabeled substrates for comparison (Figure 3.3, lanes G-J). The results of reaction time course studies (Figures 3.5) show that TES products are formed in approximately 70% yield (73 ± 1.8% and 69 ± 3.3%), regardless of which end of the substrate is radiolabeled. Because of the limitations of hand mixing relatively fast reactions, the only quantification that can be made of observed rate constants, k_{obs}, is a lower limit of >8 min^{-1} for the 3' labeled substrate (Figure 3.5A) and >6 min^{-1} for the 5' labeled substrate (Figure 3.5B) for the first step of the TES reaction. More importantly,

however, when using 3'-end radiolabeled substrate, essentially all radiolabeled intermediate is a high molecular weight band (12 ± 0.3% and 11 ± 1.6%) that corresponds to the predicted 341 nucleotide intermediate expected for the intramolecular transesterification mechanism. This is not consistent with the hydrolysis mechanism.

In addition, TES reactions were conducted with the 3'-end radiolabeled substrate, 10-mer-dG, $^{5'}$AUGACU*dG*CUC$^{3'}$, which is predicted to prevent the second step of the TES reaction (exon ligation) and should accumulate only first step product (Figure 3.4). TES reactions conducted with the 10-mer-dG substrate shows that the high molecular weight intermediate is now produced in yields of 87 ± 0.4% (Figure 3.5C). This is a further indication that the high molecular weight band is the reaction intermediate, and suggests that the TES reaction proceeds through the intramolecular transesterification mechanism. Again, the only quantification that can be made of the observed rate constant, k_{obs}, for the first step of this TES reaction is a lower limit of >6 min^{-1}.

For confirmation of the identity of TES reaction intermediates, RT-PCR was employed to amplify the high molecular weight intermediates for subsequent sequencing (Figure 3.6). For these studies a 20-mer substrate, $^{5'}$AUGACUACUCUCGU*G*CUCUU$^{3'}$, was used to facilitate RT-PCR amplification. Our results from sequencing five reaction products show that the TES reaction does produce the intermediate expected for the intramolecular transesterification reaction, whereby the 3'-end of the substrate ($^{5'}$ACUCUCGU*G*CUCUU$^{3'}$) attaches to the terminal nucleotide (G336) of the ribozyme (Figure 3.7A).

TES Reactions Using Terminal-Modified Ribozymes

A defining characteristic of the intramolecular transesterification cleavage reaction is that the ribozyme will utilize a terminal guanosine as an intramolecular nucleophile and that it must bind to the guanosine-binding site (GBS), which preferentially binds guanosine. To test if only guanosine can be utilized as the intramolecular nucleophile, PCR templates were used to synthesize two new ribozyme constructs, one ending in guanosine (rPC-ωG) and one ending in adenosine (rPC-ωA). If only a 3' terminal guanosine can be used, then only rPC-ωG will be active; rPC-ωA will produce neither product nor intermediate. TES reactions were conducted using the same

reaction conditions as for the rPC ribozyme (Figure 3.3, lanes K-R). The results show that both ribozymes produced intermediate and product with approximately the same amount of product being formed (80 ± 1.5% versus 76 ± 2.2%). The terminal adenosine, however, results in a slower reaction (k_{obs} = 0.14 ± 0.02 min^{-1}) than that with the terminal guanosine (k_{obs} = 2.6 ± 0.09 min^{-1}) (Figure 3.8). Nevertheless, the identity of the 3' terminal end of the ribozyme does affect TES reaction kinetics, as expected only for the intramolecular transesterification mechanism.

The observation that the rPC-ωA ribozyme produces both intermediate and product can be explained through two different scenarios. In the first, the non-encoded addition of nucleotides to the 3'-end of RNA transcripts could potentially lead to guanosines being added to the 3'-end of the ribozyme, which has previously been shown to be a consequence of run-off transcription utilizing T7 RNA polymerase (92, 93). For the second, a 3' terminal adenosine could potentially be used as an intramolecular nucleophile. This in fact has previously been demonstrated for a *T. thermophila* ribozyme construct that predominately terminates in an adenosine (94). The RT-PCR strategy described above was employed to isolate and sequence the TES ribozyme intermediate from TES reactions conducted with rPC-ωA ribozyme. Results show that the terminal adenosine is the nucleophile in these reactions (Figure 3.7B). Apparently, the intramolecular transesterification reaction can be catalyzed by a terminal guanosine or adenosine. That guanosine is ~20-fold more effective (Figure 3.8) is likely related to the fact that guanosine is expected to have enhanced affinity for the guanosine binding site (GBS) of the ribozyme.

Identification of TES Reaction Intermediates In Vivo

We have previously shown that TES ribozymes can be used for the excision of RNA sequences from within mRNA transcripts utilizing an *E. coli* model system (9). Briefly, a ribozyme was developed to sequence-specifically excise an internal point mutation from within a green fluorescent protein (GFP) transcript, thus repairing the mutant transcript and creating fluorescent cells (9). It was of interest, therefore, to determine if the intramolecular transesterification mechanism is occurring *in vivo*; and if so, how the ribozyme obtains the reactive 3' terminal nucleotide. To express the GFP-

specific ribozyme (termed rPC-GFP) *in vivo*, the ribozyme sequence was flanked by a T7 RNA polymerase promoter and terminator for proper initiation and termination of transcription (95). Although the terminator sequence has previously been shown to predominately end in guanosine, shown in Figure (3.9), it was not known if this guanosine was acting as an intramolecular nucleophile for the TES reaction (96).

TES reactions were conducted *in vivo* under previously described methods (9). Total RNA isolated from *in vivo* experiments were subjected to RT-PCR reactions using primers designed to bind the 5'-end of the rPC-GFP ribozyme and the 3'-end of the GFP resulting in the amplification of only the TES reaction intermediate, if present (Figure 3.10 and 3.11). A single product was amplified by RT-PCR reactions which were ligated into a cloning vector and subsequently sequenced for the presence of the TES reaction intermediate. Sequence analysis of 17 cloned RT-PCR products showed that the 3'-end of the GFP transcript was covalently attached to the 3'-end of the rPC-GFP ribozyme (Table 3.1). Additionally, in 11 of these cases, the ribozyme was covalently attached to the expected site on the target (U607; $^{5'}$-UACCUU$_{607}$-$^{3'}$) (Table 3.1). However, in 6 of these cases, the ribozyme was attached to a similar, but distinct site (U624; $^{5'}$-GCCCUU$_{624}$-$^{3'}$) (Table 3.1). Apparently, the sequence-specificity of this particular ribozyme construct is not especially high for cleavage at its intended target. Nevertheless, finding the TES ribozyme intermediate strongly suggests that the intramolecular transesterification mechanism also occurs *in vivo*.

Surprisingly, in 16 of the 17 cases the T7 terminator sequence was not present at the 3'-end of the ribozyme (Figure 3.9), which suggests that it was somehow removed *in vivo* prior to the ribozyme becoming functional (Table 3.1). For the other case, a small portion of the T7 terminator sequence was present at the 3'-end of the ribozyme. In 16 sequences the active ribozyme ended with the natural terminal nucleotide of the ribozyme (G336), and in 1 case it ended with a guanosine 4 bases downstream (G340) (Figure 3.9). It is likely that these terminal guanosines are being created through a ribozyme-mediated intramolecular 3' hydrolysis reaction, which is well known to occur *in vitro* (13, 15, 17, 68, 69, 97). If true, the ribozyme is catalyzing its own activation prior to catalyzing the TES reaction. Taken together, these results strongly support a mechanism where a

ribozyme 3' terminal guanosine is the substrate cleavage nucleophile in the TES reaction, both *in vitro* and *in vivo*.

3' Hydrolysis Can Activate the Ribozyme for the Intramolecular Transesterification Reaction

In order to test whether TES ribozymes can generate 3' terminal guanosines via 3' hydrolysis, the TES and RT-PCR procedures were conducted *in vitro* using a ribozyme construct containing an elongated 3'-end. The ribozyme, termed rPC-*Hind*III was created by linearizing the PC plasmid with *Hind*III (instead of *Xba*I), which linearizes the plasmid seven bases downstream of G336. TES reactions conducted *in vitro* using 5' and 3'-end radiolabeled 10-mer substrates resulted in the formation of the expected intermediates and products (Figure 3.3, lanes S-V), showing the construct is still active even though it was synthesized to be longer than necessary. Sequence analysis (total of five sequences) of the TES ribozyme intermediate showed that the expected intramolecular transesterification intermediate is found, including that it ends in G336. Apparently, the extra seven bases were removed *in vitro*, thus activating G336 for subsequent TES reactivity, as expected.

DISCUSSION

TES Reaction Mechanism

In this study, it was determined that the identity of the 3' terminal nucleotide of the *P. carinii* ribozyme is important for effective catalysis of the substrate cleavage reaction in TES reactions; that a 3' guanosine is required for this reactivity; that substrate cleavage is not significantly driven by ribozyme-mediated hydrolysis; that the 3' half of the target transcript is covalently attached to the 3' terminal guanosine of the ribozyme in the intermediate step of the TES reaction; and that this mechanism is occurring *in vitro* and *in vivo*. These results are compiled into the reaction mechanism shown in Figure 3.1.

These results suggest the following reaction mechanism for the first step of the TES reaction. First, the ribozyme binds the target transcript via the RE1 molecular recognition component, resulting in the formation of helix P1. The site of first step

reactivity (substrate cleavage) is thus determined as the 3'-end of helix P1, usually characterized by a G•U wobble base-pair (48, 49). The 3' terminal guanosine (in this case G336) appears to bind to the catalytic core of the ribozyme via the GBS, hence the preference for a 3' terminal guanosine, aided presumably by P9.0 helix formation. This conformationally orients the 3' hydroxyl group of the ribozyme for nucleophilic attack at the phosphodiester backbone of the substrate cleavage site. The result of this substrate cleavage reaction step is that the 3' half of the target transcript becomes covalently attached to the ribozyme. Concurrently, the 5' half of the target transcript remains bound to the ribozyme through secondary and tertiary interactions via helix P1.

The second reaction step of the TES reaction, called exon-ligation, is catalyzed by the free 3' hydroxyl group in helix P1, which is typically a uridine (48, 49). This nucleophile attacks the phosphodiester backbone upstream and adjacent to the target transcript guanosine now in the GBS, resulting in the formation of a normal phosphodiester bond between the regions flanking the excised segment. As occurs in self-splicing (46, 47, 59), a local conformation change in the ribozyme occurs, prior to exon ligation, which results in the sequential rearrangement of guanosines in the GBS. The 3' terminal guanosine of the ribozyme is replaced by the 3' guanosine of the target's excised segment. The overall reaction, then, has the excised segment being removed from the target transcript.

TES Reaction In Vivo

The results demonstrate that the intramolecular transesterification cleavage mechanism is also occurring when the TES reaction takes place in *E. coli*. Sequencing of reaction products indicate that the same reaction intermediates are generated *in vivo* as *in vitro*. Moreover, it was discovered that, in a cell, the ribozyme becomes truncated at its 3'-end, resulting in the creation of an active 3' terminal guanosine at the 3'-end of the ribozyme (Figure 3.9). It is likely that this truncation is catalyzed by the ribozyme itself, and that the mechanism is site-specific 3' hydrolysis (13, 15, 17, 68, 69, 97). In addition, sequencing of TES reaction intermediates demonstrates that the ribozyme targets two similar sites on the GFP transcript (out of hundreds of possible sites), with the desired site being targeted 65% of the time. Previous studies using the *P. carinii* ribozyme have

shown that stable base-pairing between the substrate and the first four nucleotides of the ribozymes RE1 region (to form a functional P1 helix) is primarily responsible for the molecular recognition components that bind the substrate to the ribozyme (98). It is not surprising, then, to find that both the major and minor target sites are able to form this four nucleotide helix, which accounts for their similar reactivities. Taken together, the *in vivo* reaction demonstrates that the TES is following an intramolecular transesterification cleavage mechanism, the ribozyme is self-activating, and the reaction is only moderately sequence-specific.

The 3' Terminal Adenosine Ribozyme and Role of Molecular Recognition

Results demonstrate that the *P. carinii* ribozyme can utilize a 3' terminal adenosine (A336) as the intramolecular transesterification nucleophile. Presumably, the terminal adenosine binds to the GBS of the ribozyme, which is known to preferentially bind guanosine (57). Studies using a *T. thermophila* ribozyme, however, suggest that adenosine can bind the GBS with the help of non GBS interactions (94). The most likely interaction is the formation of helix P9.0 (Figure 3.1B), which participates in the alignment of the 3' splice site of group I introns into the GBS (60-66). It seems for the TES reaction, the role of P9.0 helix formation could be multifaceted in that it helps to align the terminal nucleotide of the ribozyme into the GBS prior to the intramolecular transesterification reaction, and also aligns the substrate guanosine (the guanosine to be excised) into the GBS for the subsequent exon-ligation reaction (Figure 3.1B).

Comparison with Previous Results

That the 3' terminal guanosine of group I intron-derived ribozymes can be a nucleophile has previously been reported, although not in the context of the TES reaction. The terminal guanosine of the *T. thermophila* group I intron-derived ribozyme has been implicated in ribonuclease (15), phosphotransferase (13), and phosphatase (13) activities *in vitro*. In these reactions, the ribozymes 3' terminal guanosine was obtained through site-specific 3' hydrolysis, which exploited high temperature ribozyme pre-incubations under alkaline conditions (15, 17, 68, 69). We found that production of active rPC ribozyme *in vitro* does not require this pre-incubation, although 3' hydrolysis might still

be occurring, albeit at prevailing reaction conditions. The *T. thermophila* intron also utilizes a 3' terminal guanosine for catalyzing the reverse splicing reaction (79, 80), in which the intron becomes imbedded into cellular transcripts as well as *in vitro* cyclization reactions (67-70). Other group I ribozymes have exploited their 3' terminal guanosines for polymerization (76, 77) and recombination (78) reactions. Apparently, 3' terminal guanosines of group I intron-derived ribozymes can be involved in many natural and unnatural reactions.

Group I intron-derived ribozymes have additionally been exploited for catalyzing reactions where a 3' terminal guanosine is not required. In these instances, an exogenous guanosine cofactor is usually the first (or only) reaction step nucleophile. For example, group I intron-derived ribozymes have been used to replace terminal (81-90) segments of RNA transcripts, resulting in the repair of mutant transcripts. Apparently, group I intron-derived ribozymes can catalyze a wide variety of reactions, which demonstrates their versatility in both catalysis and molecular recognition.

Design Principles for the P. carinii Ribozyme In Vivo

Synthesizing RNAs that contain specific terminal sequences is challenging because RNA polymerases tend to add non-templated nucleotides onto the ends of RNA transcripts (92, 93). One common strategy is to use a small self-cleaving ribozyme to truncate the elongated target transcript at the desired site (72, 99). Unfortunately, this method is invasive and creates a 2', 3' cyclic phosphate, which would eliminate TES reactivity, though the 2', 3' cyclic phosphate could subsequently be removed by dephosphorylation to render the active form of the ribozyme. Fortunately, at least in *E. coli* (and confirmed *in vitro*), TES ribozymes appear to self-catalyze their own truncation, thus creating a functional 3' terminal guanosine. The efficiency with which the 3' hydrolysis reaction occurs with the rPC ribozyme, however, has not yet been determined *in vivo*. This might be a limiting factor for TES reaction conducted *in vivo*. If true, enhancing the 3' hydrolysis reaction could enhance the yield of TES ribozymes. To this end, it has been reported that P10 helix formation prior to the first step of the self-splicing reaction increases the amount of 3' hydrolysis product obtained *in vitro* (100). This

design strategy could be implemented by adding sequences after G336 that are complementary to RE3 of the ribozyme (see Figure 3.1).

In this study it was demonstrated that the ribozyme is not entirely specific for the intended site on the target. This lack of absolute sequence-specificity was found to be true with other group I intron-derived ribozyme-mediated reactions (82). Therefore, a variety of strategies has been devised to ameliorate this limitation. Chief among them is a clever strategy to add an artificial substrate binding element to a trans-splicing ribozyme (83, 84, 86-90). This extended guide sequence (EGS) could easily be inserted at the 5'-end of the ribozyme to help increase sequence specificity of TES ribozymes.

Prior to this study, no reasonable experimental strategies could be used to determine sites within an RNA transcript that are accessible to the rPC ribozyme. Our results, however, suggest that accessible sites can be determined by amplifying the intermediates using RT-PCR. Intermediates will only be formed if their target sites are accessible. This strategy could be implemented by using one primer that is specific for the ribozyme and the other primer can be specific for the 3'-end of the transcript.

Implications for Native Group I Intron Function In Vivo

Group I introns catalyze the self-splicing reaction, wherein the introns excise themselves (sometimes with the requirement of proteins), from RNA transcripts. It has also been demonstrated that intact group I introns can be mobile elements by catalyzing a reverse-splicing reaction (79, 80). We now demonstrate that essentially intact group I introns can covalently attach themselves to RNA transcripts *in trans* while catalyzing the removal of an internal segment from within RNA transcripts through TES. Apparently, group I introns are multifaceted *in vivo* catalysts.

Table 3. 1. Analysis of ribozyme intermediates isolated from *in vivo* testing from JM109(DE3) cells

Test construct	RT-PCR product[a]	TES intermediate utilizing G336[b]	TES intermediate utilizing G340[c]	Primary site targeted[d]	Alternative site targeted[e]
Active ribozyme + Mutant GFP	+	94%	6%	65%	35%

[a] Total RNA isolated from JM109 (DE3) cells were subjected to RT-PCR to amplify the ribozyme intermediate with the 3'-end of the GFP mRNA attached to the ribozyme; (+) positive RT-PCR band. [b] Percentage of ribozyme intermediates with 3'-end of GFP mRNA attached through G336 of the GFP-rPC ribozyme (16 of 17 RT-PCR products sequenced). [c] Percentage of ribozyme intermediates with 3'-end of GFP mRNA attached through G340 of the GFP-rPC ribozyme (1 of 17 RT-PCR products sequenced). [d] Percentage of RT-PCR products sequenced (11 of 17 RT-PCR products) that indicated the GFP-rPC ribozyme targeted the primary site (U607, Codon 201). [e] Percentage of RT-PCR products sequenced (6 of 17 RT-PCR products) that indicated the GFP-rPC ribozyme targeted an alternative site (U624, Codon 207).

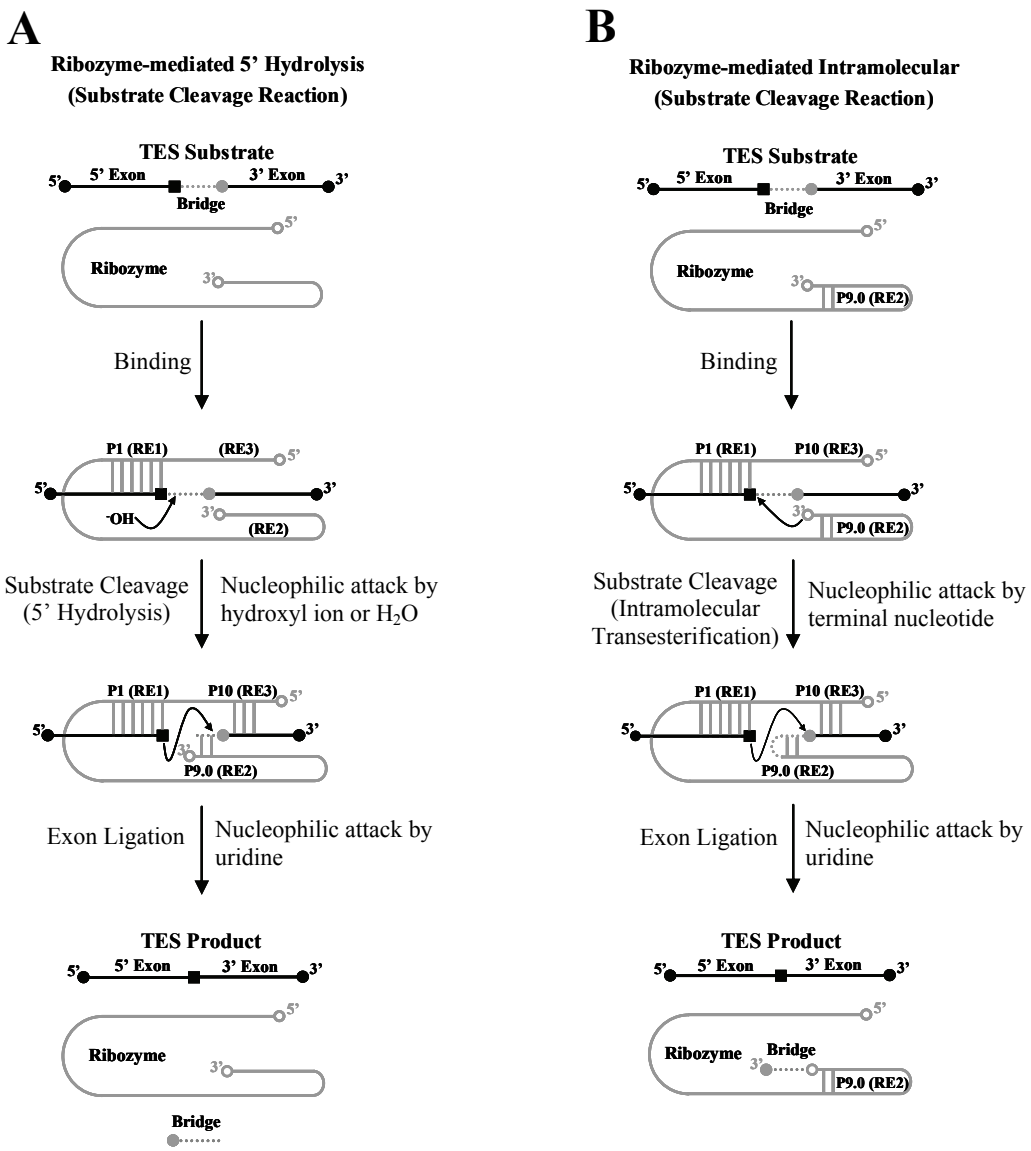


Figure 3.1. Comparison of potential TES reaction mechanisms. The *P. carinii* ribozyme is shown as a grey line and the 5' and 3' exons are black lines. The bridge region (the region to be excised) is shown as a dashed gray line. The black square within the 5' exon represents a uridine and the gray circle adjacent to the 3' exon represents a guanosine. The open circle (B) at the 3'-end of the ribozyme represents the 3' terminal nucleotide of the ribozyme (ωG or ωA). The first step of the TES reaction proceeds through either ribozyme-mediated 5' hydrolysis (A) or ribozyme-mediated intramolecular transesterification (B).

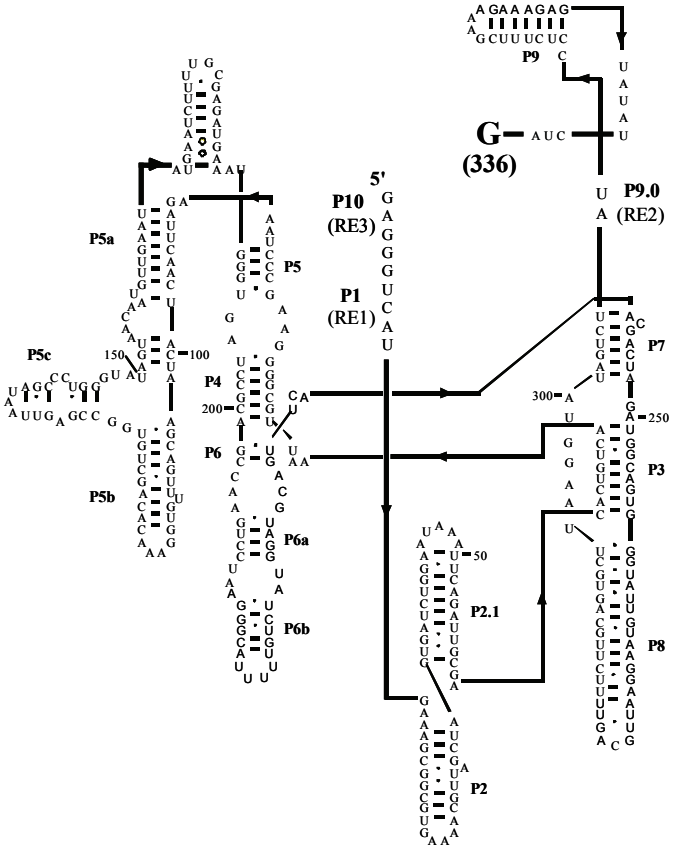


Figure 3.2. Proposed secondary structure of the *Pneumocystis carinii* (rPC) ribozyme. The rPC ribozyme is denoted in uppercase lettering. The molecular recognition elements are denoted in bold uppercase. The potential 3' terminal guanosine, G336, of the ribozyme is in bold lettering.

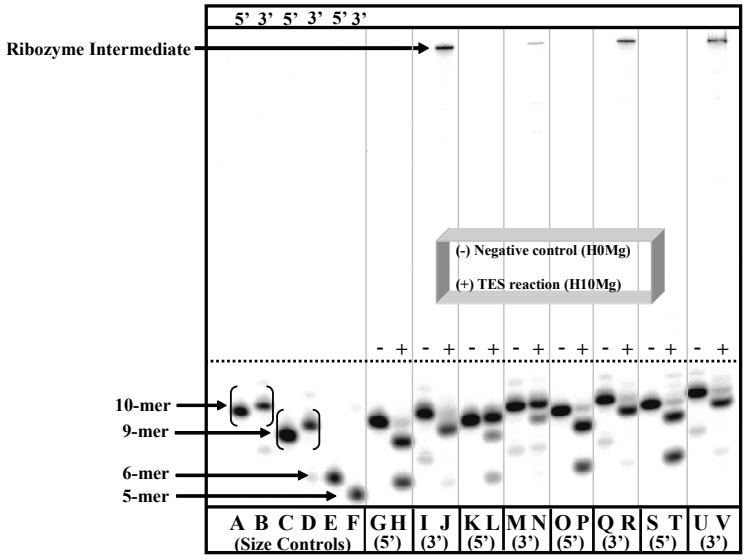


Figure 3.3. Representative polyacrylamide gel for TES conducted with 5' and 3'-end radiolabeled substrates. Polyacrylamide gel showing the substrates, intermediates, and products of TES reactions conducted with 166 nM ribozyme and 1.33 nM substrate at 44 °C for 15 min. Lanes A, C, and E contain 5'-end radiolabeled 10-mer, 9-mer, and 6-mer size controls. Lanes B, D, and F contain 3'-end radiolabeled 10-mer, 9-mer, and 5-mer size controls. Note that the 3'-end radiolabeled size controls are one nucleotide larger than 5'-end radiolabeled size controls. Gel lanes containing a plus (+) sign show reactions conducted in H10Mg and lanes containing a minus (-) sign show reactions conducted in H0Mg. Lanes G-J contain the normal rPC ribozyme with 5'-end radiolabeled 10-mer (lanes G and H) or 3'-end radiolabeled 10-mer (lanes I and J). Lanes K-N contain the rPC-ωA ribozyme with 5'-end radiolabeled 10-mer (lanes K and L) or 3'-end radiolabeled 10-mer (lanes M and N). Lanes O-R contain the rPC-ωG ribozyme with 5'-end radiolabeled 10-mer (lanes O and P) or 3'-end radiolabeled 10-mer (lanes Q and R). Lanes S-V contain the rPC-*Hind*III ribozyme with 5'-end radiolabeled 10-mer (lanes S and T) or 3'-end radiolabeled 10-mer (lanes U and V).

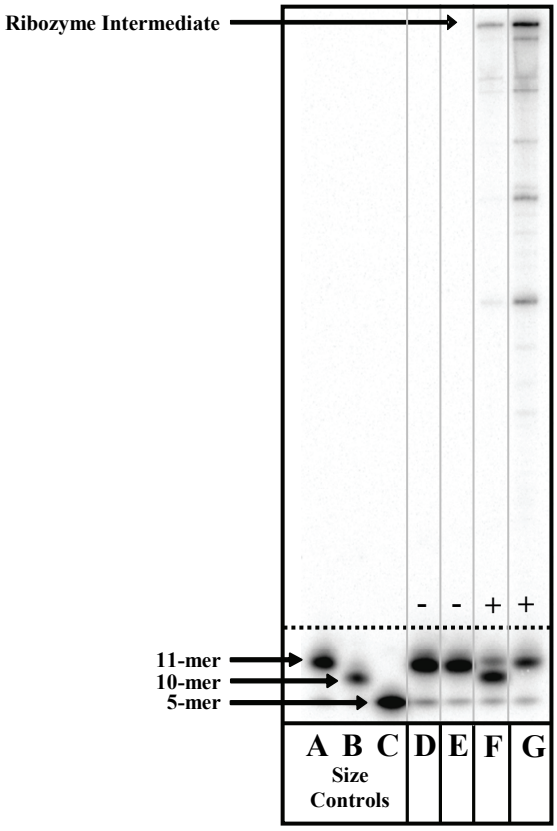


Figure 3.4. Representative polyacrylamide gel for TES reactions conducted with the 10-mer-dG substrate. Polyacrylamide gel showing TES substrate, intermediate, and product of TES reactions conducted with 166 nM ribozyme, 1.33 nM substrate, H10Mg, at 44 °C for 15 min. Lanes A, B, and C contain 3'-end radiolabeled 11-mer, 10-mer, and 5-mer size controls, respectively. Note that 3'-end radiolabeled size controls are one nucleotide larger than 5'-end radiolabeled size controls. Plus (+) represent TES reactions conducted in H10Mg and minus (-) represent TES reactions conducted in H0Mg (Lane D) or without ribozyme (Lane E). Lane F contained the normal rPC ribozyme with 3'-end radiolabeled 10-mer. Lane G contains the normal rPC ribozyme with the 3'-end radiolabeled 10-mer-dG substrate.

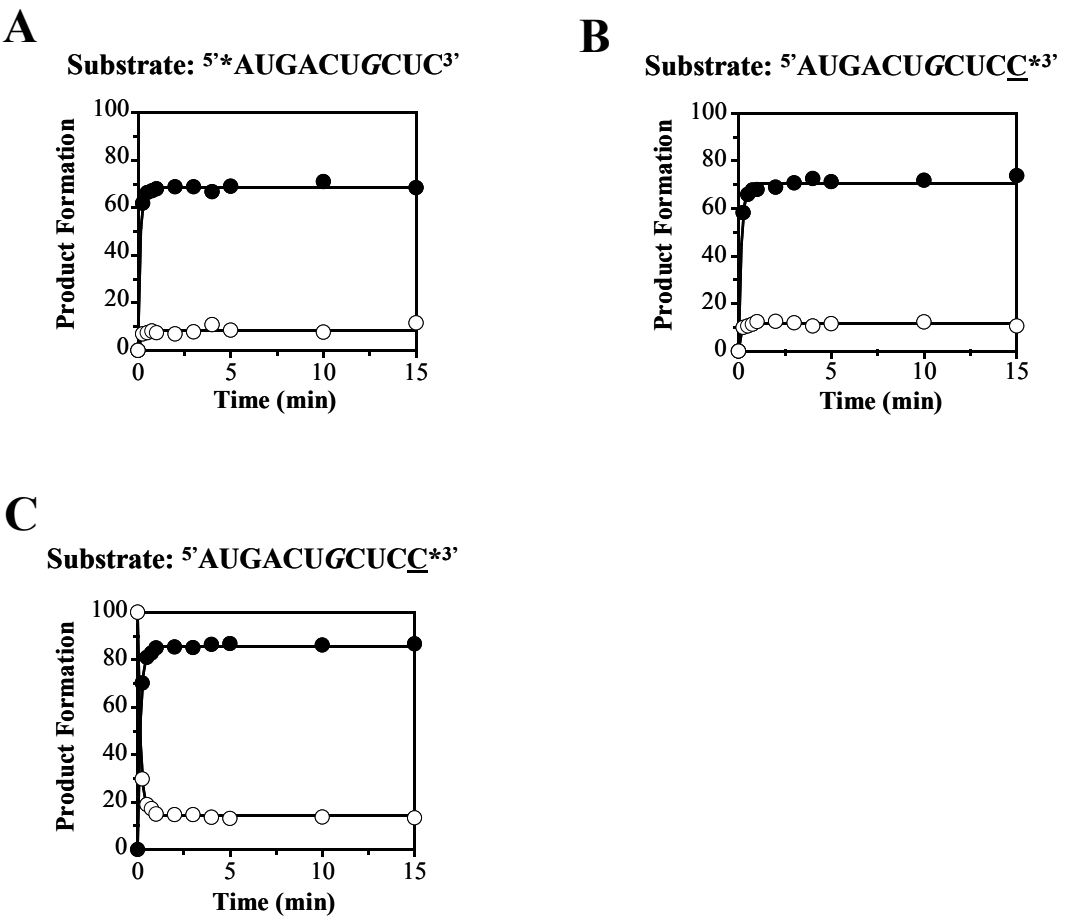


Figure 3.5. Graphical representations of TES reactions as a function of substrate. Reactions were conducted in duplicate with 166 nM rPC, 1.33 nM substrate, and H10Mg at 44 °C for 15 min. A. TES reaction using 5'-end radiolabeled substrate (AUGACU*G*CUCC). The 10-mer TES product (k_{obs} >6 min^{-1}) is represented by filled circles, and the higher molecular weight (ribozyme) intermediate is represented by open circles. B. TES reaction using 3'-end radiolabeled substrate (AUGACU*G*CUC). The 9-mer TES product (k_{obs} >8 min^{-1}) is represented by filled circles and the 6-mer intermediate is represented as open circles. C. TES reaction using 3'-end radiolabeled substrate (AUGACU*dG*CUCC). The 10-mer TES product is represented by filled circles, and the higher molecular weight (ribozyme) intermediate (k_{obs} >6 min^{-1}) is represented by open circles. Each graph shows the average of two independent assays. Standard deviations for each time point are less than 15%.

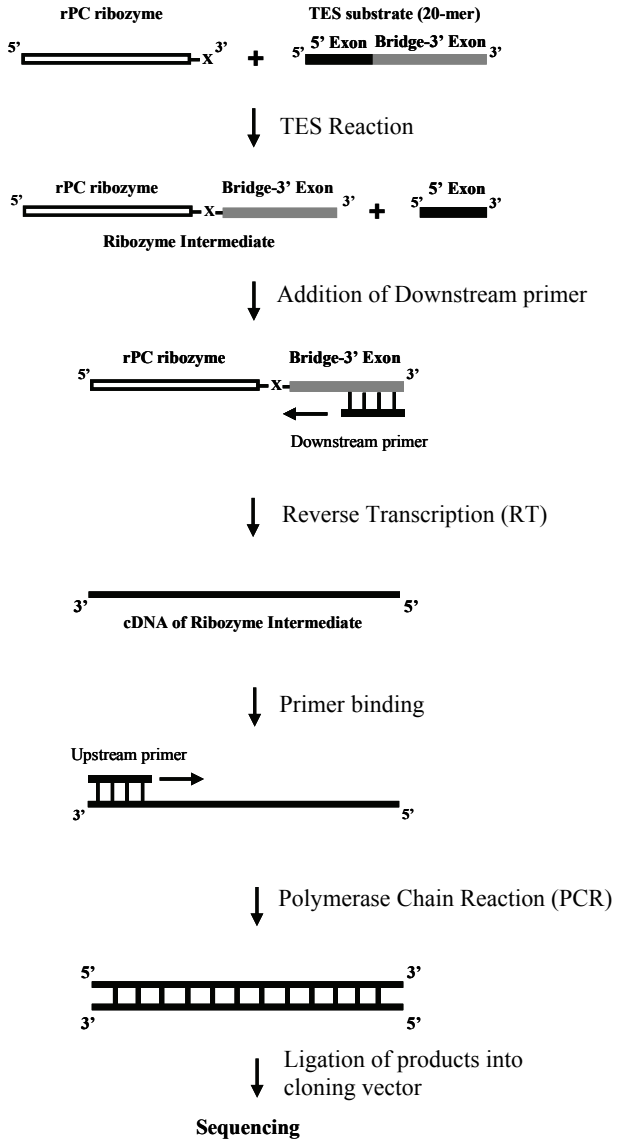
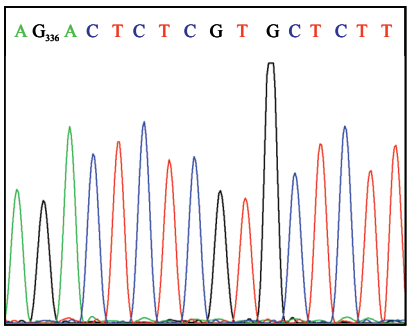


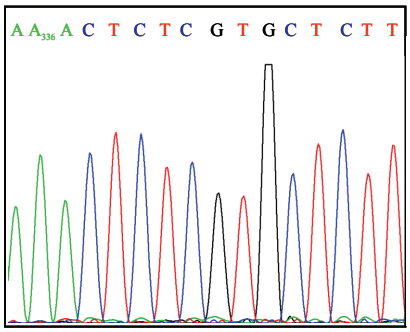
Figure 3.6. Diagram of RT-PCR strategy for the amplification of TES ribozyme intermediate. TES reactions conducted with rPC and rPC variants using the 20-mer TES substrate ($^{5'}$AUGACUACUCUCGU*G*CUCUU$^{3'}$). Reaction with the substrate will result in the covalent attachment of the bridge-3' exon region ($^{5'}$ACUCUCGU*G*CUCUU$^{3'}$, denoted as grey block) to the 3'-end of the rPC ribozyme. The bridge-3' exon region is used as a primer binding site for reverse transcription (RT) for first strand synthesis of the TES ribozyme intermediate. An additional primer, specific for the 5'-end of the rPC ribozyme, is included for the polymerase chain reaction (PCR) to amplify the TES

ribozyme intermediate, if present. Isolated RT-PCR products are subsequently ligated into a cloning vector and sequenced for identification of TES ribozyme intermediates.

A



B



C

Table 1. Sequencing results for RT-PCR products isolated from TES reactions conducted with rPC and rPC-ωA ribozymes.

Ribozyme	Number of sequences (RT-PCR products)	Identity of 3' terminal nucleotide[a]	Sequence of ribozyme+ bridge-3'exon[b]
rPC	5	G336	$^{5'}$AG$_{336}$ACUCUCGUGCUCUU$^{3'}$
rPC-ωA	5	A336	$^{5'}$AA$_{336}$ACUCUCGUGCUCUU$^{3'}$

[a] RT-PCR products were isolated from TES reactions conducted with either the rPC or rPC-A ribozyme construct. Sequencing demonstrated that the 3' terminal nucleotide was either a guanosine (G336 for rPC) or adenosine (A336 for rPC-ωA). [b] Sequencing of 5 RT-PCR products demonstrated that the bridge-3' exon region of the 20-mer substrate ($^{5'}$AUGACUACUCUCGUGCUCUU$^{3'}$) became covalently attached to the 3' terminal nucleotide of each ribozyme.

Figure 3.7. Sequencing results for RT-PCR products isolated from TES ribozyme intermediates. A. Representative chromatogram for sequencing result for the rPC ribozyme. Note that the 3'-end of the substrate attaches to the terminal nucleotide (G336) of the ribozyme. Individual nucleotides are denoted by color; guanosine (black), adenosine (green), cytosine (blue), and thymine (red). B. Representative chromatogram for sequencing result for the rPC-ωA ribozyme. The 3' terminal nucleotide (A336) is depicted in green. Note that the 3'-end of the substrate attaches to the terminal nucleotide (A336) of the ribozyme. C. Table summarizing all five sequencing results for the rPC and rPC-ωA ribozymes.

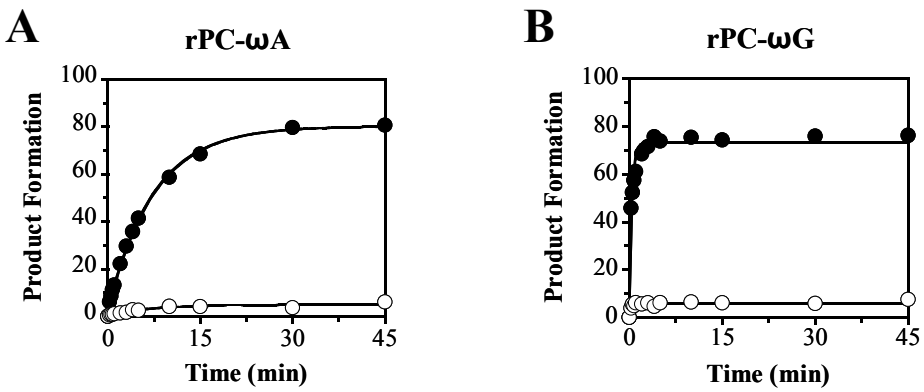


Figure 3.8. Graphical representations of TES reactions as a function of ribozyme construct. Reactions were conducted in duplicate with 166 nM ribozyme, 1.33 nM substrate, and H10Mg at 44 °C for 15 min. A. TES reaction using the PCR-derived rPC-ωA ribozyme construct. The 10-mer TES product (k_{obs} = 0.14 ± 0.02 min^{-1}) is represented by filled circles, and the higher molecular weight (ribozyme) intermediate is represented by open circles. B. TES reaction using the PCR-derived rPC-ωG ribozyme construct. The 10-mer TES product (k_{obs} = 2.6 ± 0.09 min^{-1}) is represented by filled circles, and the higher molecular weight (ribozyme) intermediate is represented by open circles. Each graph shows the average of two independent assays. Standard deviations for each time point are less than 15%.

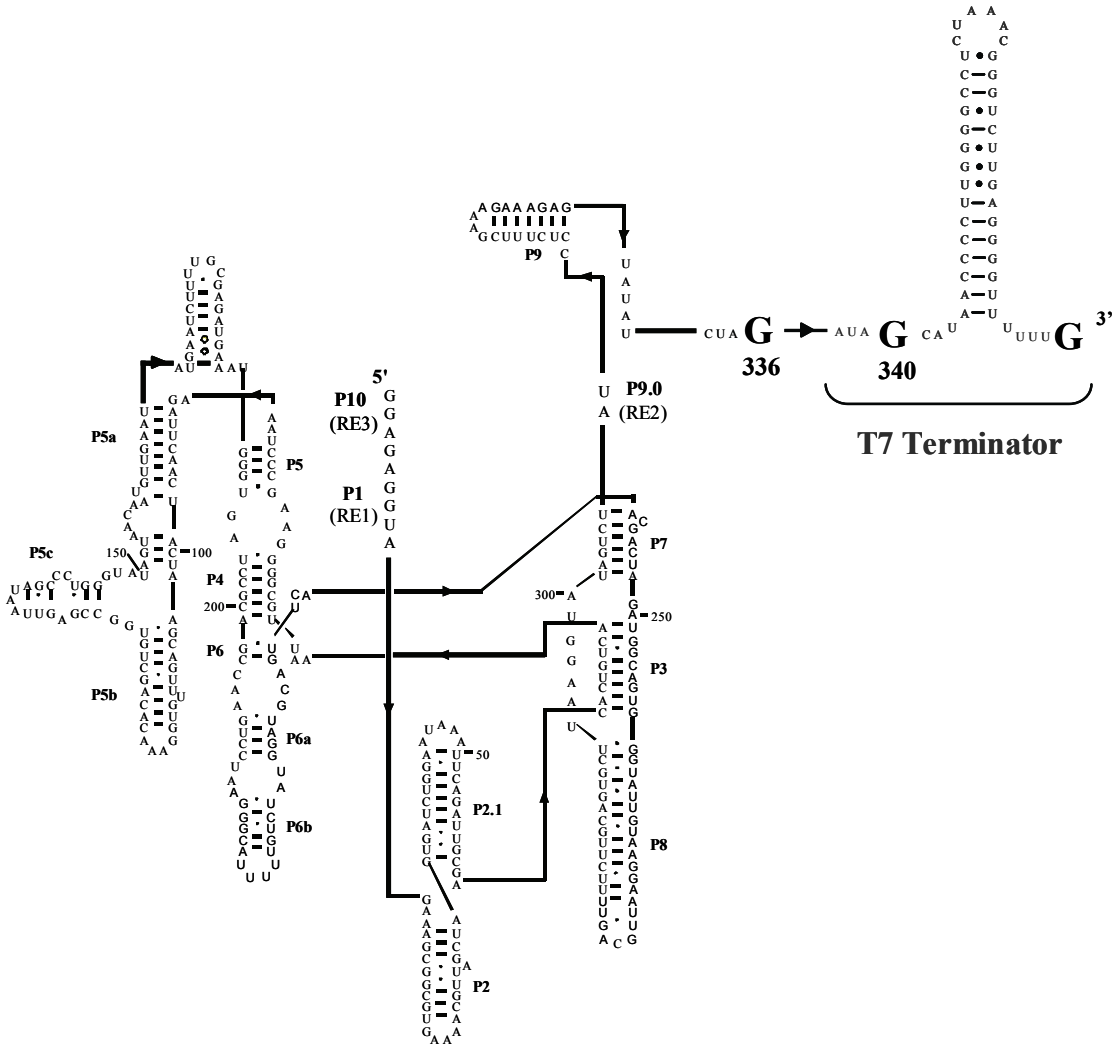


Figure 3.9. Sequence and proposed secondary structure of the rPC-GFP ribozyme. The previously constructed rPC-GFP ribozyme (8) designed for *in vivo* model test system is shown in uppercase lettering. The sequence and predicted structure of the T7 terminator, which was appended onto the 3'-end of the rPC-GFP ribozyme for *in vivo* experiments are highlighted. The guanosines (G336 and G340) found to act as ribozyme 3' terminal guanosines are denoted.

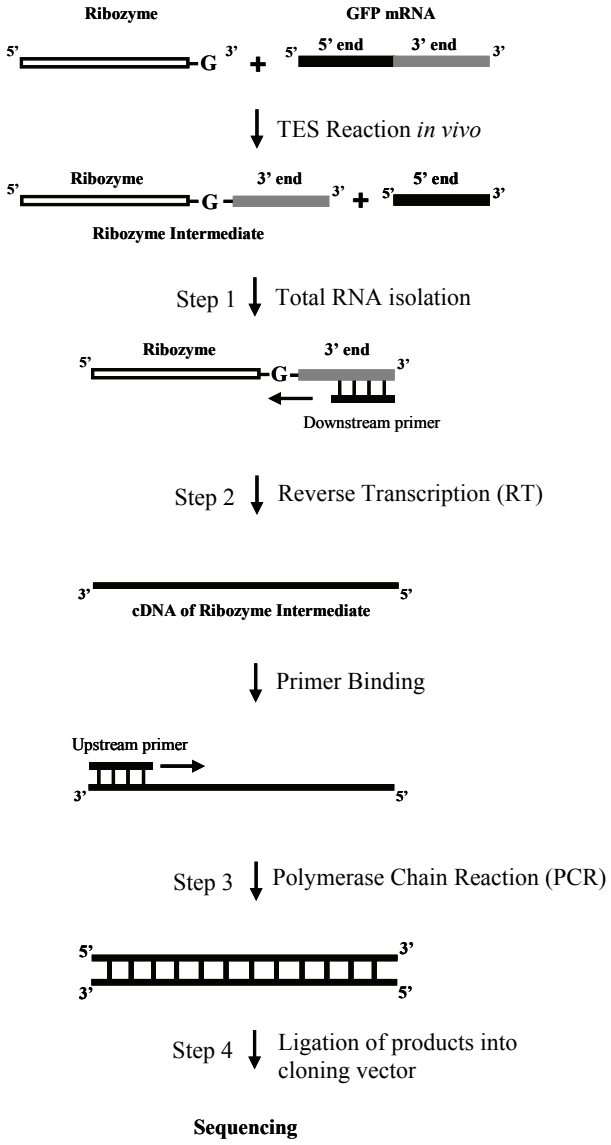
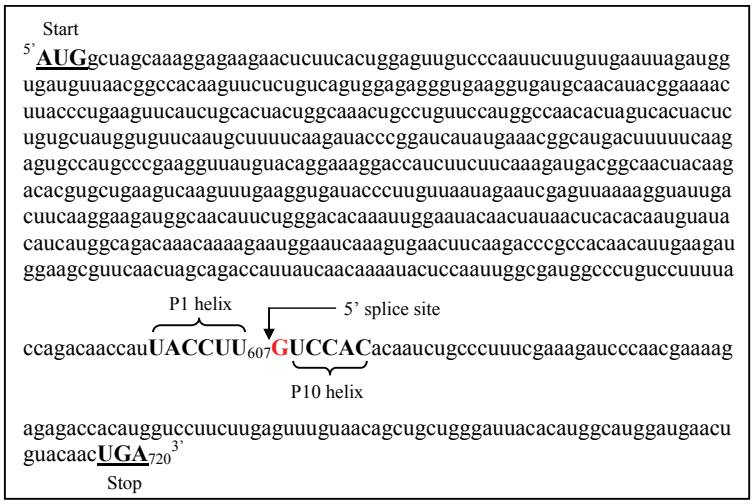


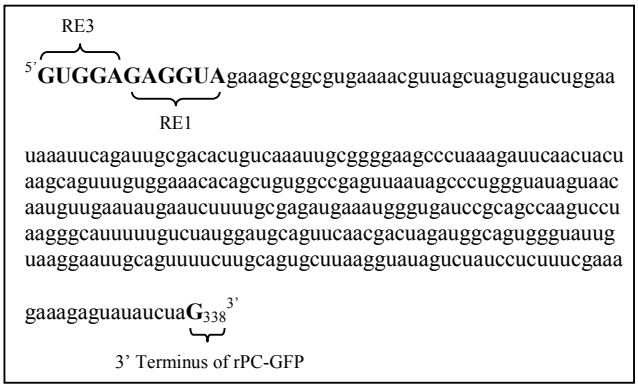
Figure 3.10. RT-PCR strategy for isolation of the TES intermediate *in vivo*. The endogenous guanosine-mediated 5' cleavage reaction *in vivo* will result in the covalent attachment of the 3'-end of the GFP mRNA to the 3'-end of the *P. carinii* ribozyme (GFP-rPC) through a 3' terminal guanosine. The TES ribozyme intermediate, if present, will constitute a portion of the total RNA isolated from JM109 (DE3) cells (Step 1). In step 2, the 3'-end of the GFP transcript is used as a primer binding site for reverse transcription (RT) for first strand synthesis of the TES ribozyme intermediate. An additional primer, specific for the 5'-end of the GFP-rPC ribozyme, is used for polymerase chain reaction (PCR) in step 3 to amplify the TES ribozyme intermediate.

Isolated RT-PCR products are subsequently ligated into a cloning vector (Step 4) and sequenced for identification of TES ribozyme intermediates.

A


Start
5' AUGgcuagcaaaggagaagaacucuucacuggaguugucccaauucuuguugaauuagaugg
ugauguuaacggccacaaguucucugucaguggagagggugaaggugaugcaacauacggaaaac
uuacccugaaguucaucugcacuacuggcaaacugccuguuccauggccaacacuagucacuacuc
ugugcuaugguguucaaugcuuuucaagauacccggaucauaugaaacggcaugacuuuuucaag
agugccaugcccgaagguuauguacaggaaaggaccaucuucuucaaagaugacggcaacuacaag
acacgugcugaagucaaguuugaaggugauacccuuguuaauagaaucgaguuaaaagguauuga
cuucaaggaagauggcaacauucugggacacaaauuggaauacaacuauaacucacacaauguaua
caucauggcagacaaacaaaagaauggaaucaaagugaacuucaagacccgccacaacauugaagau
ggaagcguucaacuagcagaccauuaucaacaaaauacuccaauuggcgauggcccuguccuuuua
P1 helix
5' splice site
ccagacaaccauUACCUU607GUCCACacaaucugcccuuucgaaagaucccaacgaaaag
P10 helix
agagaccacaugguccuucuugaguuuguaacagcugcugggauuacacauggcauggaugaacu
guacaacUGA720 3'
Stop

B


RE3
5' GUGGAGAGGUAgaaagcggcgugaaaacguuagcuagugaucuggaa
RE1
uaaauucagauugcgacacugucaaauugcggggaagcccuaaagauucaacuacu
aagcaguuuguggaaacacagcuguggccgaguuaauagcccuggguauaguaac
aauguugaauaugaaucuuuugcgagaugaaaugggugauccgcagccaaguccu
aagggcauuuuuugucuauggaugcaguucaacgacuagauggcaguggguauug
uaaggaauugcaguuuucuugcagugcuuaaggguauagucuaucccucuuucgaaa
gaaagaguauaucuaG338 3'
3' Terminus of rPC-GFP

C

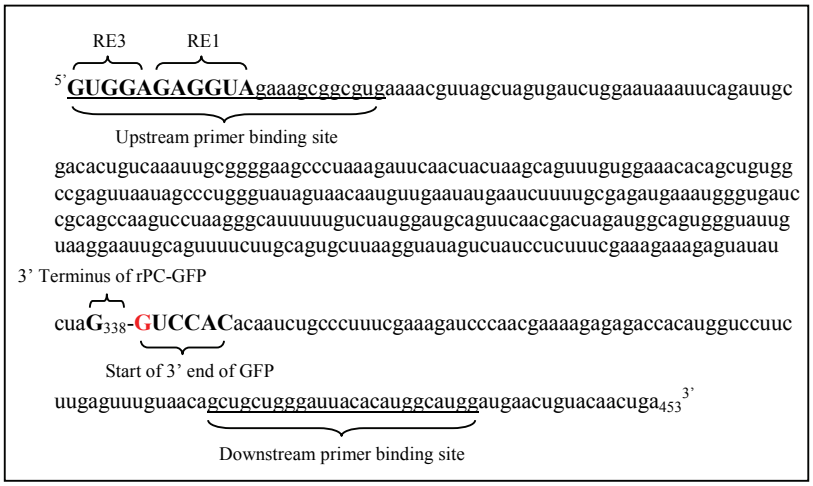

RE3
RE1
5' GUGGAGAGGUAgaaagcggcgugaaaacguuagcuagugaucuggaauaaaauucagauugc
Upstream primer binding site
gacacugucaaauugcggggaagcccuaaagauucaacuacuaagcaguuuguggaaacacagcugugg
ccgaguuaauagcccuggguauaguaacaauguugaauaugaaucuuuugcgagaugaaaugggugauc
cgcagccaaguccuaagggcauuuuuugucuauggaugcaguucaacgacuagauggcaguggguauug
uaaggaauugcaguuuucuugcagugcuuaaggguauagucuaucccucuuucgaaagaaagaguauau
3' Terminus of rPC-GFP
cuaG338-GUCCACacaaucugcccuuucgaaagaucccaacgaaaagagagaccacauggcccuuc
Start of 3' end of GFP
uugaguuuguaacagcugcugggauuacacauggcauggaugaacuguacaacuga453 3'
Downstream primer binding site

Figure 3.11. Sequences of predicted RNA transcripts for isolation of *in vivo* TES ribozyme intermediate. A. Sequence of the coding region of the GFP mRNA transcript (720 nt). The nucleotides involved in P1 and P10 helix formation are denoted in bold, black lettering. The region to be excised is denoted in bold, red lettering. The 5' splice site is denoted with an arrow. B. Sequence of the rPC-GFP ribozyme. The molecular recognition elements RE1 and RE3 and the 3' terminal guanosine (G338) are denoted in bold, black lettering. Note that the rPC-GFP ribozyme is two nucleotides longer than the normal rPC ribozyme due to RE3 consisting of five nucleotides. C. Sequence of the predicted *in vivo* TES ribozyme intermediate. The TES ribozyme product will consist of the 3'-end of the GFP transcript being covalently attached to the 3'-end of the rPC-GFP ribozyme. The TES ribozyme intermediate is predicted to be 453 nt in length. The upstream and downstream primer binding sites utilized for the RT-PCR strategy are underlined.

CHAPTER FOUR-KINETIC CHARACTERIZATION OF THE FIRST STEP OF THE RIBOZYME-CATALYZED TRANS EXCISION-SPLICING REACTION

INTRODUCTION

As previously discussed, a group I intron-derived ribozyme from *Pneumocystis carinii* can catalyze the excision of a targeted sequence from within an RNA substrate (8). The reaction, termed trans excision-splicing (TES), consists of two reaction steps; substrate cleavage (intramolecular transesterification) followed by exon ligation (Figure 4.1). In the substrate cleavage reaction, the phosphodiester backbone of TES substrates are cleaved through nucleophilic attack of a 3' terminal guanosine (G336), generating both 5' and 3' exon intermediates (see Chapter 3). In the exon-ligation step, a nucleophilic uridine (at the 3'-end of the 5' exon intermediate) performs a nucleophilic attack upon a phosphodiester backbone position within the 3' exon intermediate, resulting in ligation of the exon sequences and excision of the internal segment (Figure 4.1). The substrate cleavage reaction step is analogous to that in self-splicing (3), except that self-splicing utilizes an intermolecular nucleophile (exogenous guanosine), while the TES reaction utilizes an intramolecular nucleophile (endogenous guanosine). The TES substrate cleavage reaction is similar to the cyclization reactions performed by group I introns for generating both full-length and truncated circular introns (67-70, 101).

The establishment of kinetic frameworks has been useful for the dissection of the individual steps of RNA catalyzed reactions (16, 72, 73, 102-111). The results obtained from these types of studies have been mechanistically informative and have advanced our fundamental understanding of the chemical basis for RNA catalysis. To further investigate the fine details regarding the mechanism of the TES substrate cleavage reaction, a minimal kinetic framework for the substrate cleavage reaction was established (Figure 4.2).

There are several conclusions drawn from this kinetic framework as they relate to the TES reaction. The rate constant for the substrate cleavage reaction is approximately 60-fold lower than that reported for the self-splicing reaction using a *Tetrahymena thermophila* ribozyme, regardless of whether an intermolecular or intramolecular

guanosine is being utilized as the substrate cleavage nucleophile (16, 17). The rate constant for substrate dissociation is only 4-fold lower than that for substrate cleavage. In addition, product dissociation is slower than the rate of cleavage which suggests that multiple turnover, with respect to substrate cleavage, will likely be curtailed. The results indicate that a conformational change exists between the two steps of the TES reaction. Taken together, these results have greatly increased our current understanding of the mechanism for the TES reaction. The kinetic framework presented herein will likely lead to developing TES ribozymes that are more efficient at catalyzing the TES reaction either *in vitro* or *in vivo*.

MATERIALS AND METHODS

Oligonucleotide Synthesis and Purification

RNA oligonucleotides were obtained from Dharmacon Research, Inc. (Lafayette, CO) and were deprotected using the manufacturer's standard protocol. Oligonucleotide concentrations were calculated from UV absorption measurements using a Beckman DU 650 spectrophotometer (Beckman Coulter, Inc.; Fullerton, CA). Oligonucleotides were 5'-end radiolabeled via phosphorylation of the 5' terminal hydroxyl group with [γ-^{32}P] ATP (Amersham Pharmacia Biotech; Piscataway, NJ) using T4 polynucleotide kinase (New England Biolabs; Beverly, MA), as previously described (75).

Ribozyme Synthesis

The *Pneumocystis carinii* ribozyme plasmid, PC, was linearized in a 50 μL reaction consisting of 16 μg of plasmid, 1 X REACT 2 buffer, and 50 units *XbaI* (Invitrogen; Grand Island, NY) at 37 °C for 2 h. The linearized plasmid was isolated using the QIAquick PCR purification kit (Qiagen; Valencia, CA). Run-off transcription was performed for 2 h in a 50 μL reaction consisting of 1 μg of linear DNA, 50 units of T7 RNA polymerase (New England Biolabs), 40 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 5 mM DTT, 5 mM spermidine, and 1 mM rNTP mix. The individual ribozyme were purified by a Qiagen Plasmid Midi Kit as previously described (8).

Determination of Observed Substrate Cleavage Rate Constants (k_{obs} and k_2)

The first order rate constant for substrate cleavage, k_{obs}, was measured under single-turnover conditions, in which case the release of product would not affect the observed rate constants. All reactions were conducted at 44 °C in H10Mg buffer, which consists of 50 mM HEPES (25 mM Na^+), 135 mM KCl, and 10 mM $MgCl_2$ at pH 7.5. These reaction conditions appear to be optimal for the TES reaction (8). For the pH dependence studies, HEPES (pH 7.5) was replaced with MES (pH 5.0 - 6.8), HEPES (pH 6.8 - 7.5), or EPPS (pH 7.5 - 8.5). Reactions were initiated by adding 5 μL of an 8 nM solution of 5'-end radiolabeled substrate [r($^{5'}$AUGACU*dG*CUC$^{3'}$)] in the appropriate buffer to a 25 μL solution of various concentrations of ribozyme (6 nM-360 nM) in the same buffer. Note that the ribozyme solution was preincubated at 60 °C for 5 minutes and then allowed to slow cool to 44 °C) to facilitate folding of the ribozyme prior to the addition of the radiolabeled substrate. Aliquots (3 μL) were removed at specified times and quenched with an equal volume of 2x stop buffer (10 mM urea, 0.1x TBE, 3 mM EDTA). The substrate and products were denatured at 90 °C for 1 min and then separated on a 12.5% denaturing polyacrylamide gel. The bands were visualized on a Molecular Dynamics Storm 860 Phosphorimager and quantified using Imagequant software (Molecular Dynamics). Data were fit using the Kaleidagraph curve-fitting program (Synergy Software; Reading, PA). The final concentration of the radiolabeled substrate in all reactions is 1.3 nM. A typical reaction utilized H10Mg buffer and a final ribozyme concentration of 166 nM. Pseudo-first order rate constants for the appearance of products were fit using the following single exponential equation:

$$[P]_t = [P]_\infty(1-e^{-kt}) \quad [1]$$

In this equation $[P]_t$ and $[P]_\infty$ are the percentages of product formed at time *t* and at the end point, respectively; and k is the first order rate constant.

Determination of the Substrate Dissociation Rate Constant (k_{-1})

Pulse-chase experiments (16, 112) were used to measure the rate constant for substrate dissociation, k_{-1}. In these experiments, 10 μL of 200 nM ribozyme in H10Mg buffer was combined with 2 μL of 8 nM 5' radiolabeled substrate in H10Mg buffer for a period of t_1 = 30 s. The ribozyme solution was preincubated at 60 °C for 5 min and then

slow cooled to 44 °C before the addition of the substrate, which was also at 44 °C. The chase phase was then initiated by taking out 5 μL of the reaction mixture and diluting the reaction mixture with 25 μL of H10Mg buffer so that [E] < K_M. During the chase period, t_2, dissociation of labeled substrate from the ribozyme is essentially irreversible. Aliquots were removed at various times during the chase phase and the reaction was quenched by adding an equal volume of 2X stop buffer. An otherwise identical reaction, but without adding the chase (which in this case is buffer), was carried out in parallel. The first-order observed rate constants k_{obs}, chase and k_{obs}, no-chase were obtained from a single-exponential fit of this data using Equation 1 (as a function of t_2). The observed rate constant for substrate dissociation (k_{-1}) was then calculated (Equation 2) as the difference between the two measured observed rate constants:

$$k_{-1} = k_{obs,\ chase} - k_{obs,\ no\text{-}chase} \quad [2]$$

Determination of the Substrate Association Rate Constant (k_1)

The rate constant for substrate binding, k_1, was measured using a series of pulse-chase experiments. In each reaction, 5 μL of a ribozyme stock (from 36 to 240 nM) in H10Mg buffer was combined with 1 μL 8 nM 5'-end labeled substrate and allowed to react in a total volume of 6 μL. For each ribozyme concentration, several chase reactions were initiated. In each chase, 1 μL of the original reaction mixture was removed and diluted five-fold at various times, t_1, ranging from 15 s to 120 s. The addition of chase renders the dissociation of the substrate essentially irreversible. The chase reaction, t_2, was then allowed to proceed for 15 min, at which point the 5' cleavage reaction was essentially complete. The reaction was quenched with an equal volume of 2X stop buffer. The percent product formed during the chase period was plotted against time t_1. Observed rate constants (k_{obs}) were obtained by fitting the data to Equation 1. This observed rate constant measures the rate of approach to equilibrium where substrate association is equal to substrate dissociation. Hence, the rate of substrate association was obtained (16, 108) by plotting k_{obs} against ribozyme concentration and fitting to the equation:

$$k_{obs} = k_1\,[E] + k_{-1}$$

Determination of the Dissociation Constant of the Ribozyme-Product Complex, (K_d^P)

The equilibrium dissociation constant K_d^P of the 5' exon mimic binding to the ribozyme was determined using native polyacrylamide gel electrophoresis (75, 98, 103, 107). In this assay, several concentrations of ribozyme, ranging from 1.5 nM to 300 nM, were preannealed in 5 μL total volume containing 3.4% glycerol and H10Mg buffer for 5 min at 60 °C. After the solutions slow cooled to 44 °C, 2.5 μL of a stock of 0.5 nM radiolabeled 5' exon mimic in H10Mg buffer at 44 °C was added. The mixture was incubated at 44 °C for at least 90 min. To maintain the integrity of the bound species during gel electrophoresis, the gel and the running buffer was made of H10Mg buffer and they were prewarmed to 44 °C before the samples were loaded. The bound and unbound 5' exon mimics were separated from each other by running 6 μL of each reaction on a 10% nondenaturing polyacrylamide gel. The gel was placed on chromatography paper (Whatman 3MM CHR) and dried under vacuum for 30 min at 70 °C. The bands were visualized on a Molecular Dynamics Storm 860 Phosphorimager. The bands were quantified using ImageQuant software (Molecular Dynamics). Data were fit with Kaleidagraph curve-fitting program (Synergy Software) using the equation:

$$\theta = [\text{ribozyme}]_u / ([\text{ribozyme}]_u + K_d)$$

In this equation, K_d is the equilibrium dissociation constant of the 5' exon mimic, θ is the fraction of 5' exon mimic bound to the ribozyme, and $[\text{ribozyme}]_u$ is the concentration of unbound ribozyme in the reaction (75, 113).

Determination of Rate Constant of Sustrate Cleavage Product Dissociation (k_{-3})

The dissociation rate constant of the 5' exon intermediate (k_{-3}), was measured by a pulse-chase protocol, followed by analysis of the ribozyme/product complex using native polyacrylamide gel electrophoresis. In a typical experiment to measure k_{-3}, A solution of 300 nM ribozyme in 10 μL H10Mg buffer containing 3.4% glycerol was preincubated for 5 min at 60 °C and then allowed to slow cool to 44 °C. Then 5 μL of 0.5 nM 5'-end labeled 5' exon intermediate was added and the reaction mixture was incubated at 44 °C for 30 min to allow complete binding. A chase reaction was then initiated by the addition of 40 μL of 5.4 μM unlabeled 5' exon intermediate (in reaction buffer) to follow the dissociation of 5' exon intermediate from the ribozyme-5' exon complex. The final

concentrations of the reactants in the chase reaction were 40 nM ribozyme, 33 pmol 5'-end labeled substrate cleavage product, and 4 μM unlabeled 5' exon intermediate (as chase) in 50 μL reaction volume. Time points were taken by removing 5 μL aliquot from the reaction mixture and immediately loaded onto a running 10% native polyacrylamide gel. The dissociation rate was obtained using equation 1. Additionally, the rate constant of substrate cleavage product association (k_3) was obtained (111) using the equation:

$$k_3 = k_{-3}/K_d^P.$$

RESULTS

A kinetic scheme for the substrate cleavage reaction is summarized in Figure 4.2. One complication in studying the substrate cleavage reaction step, however, is that the second reaction step of TES (exon ligation) proceeds immediately after the first step (19). Therefore, a previously described modified substrate system [r($^{5'}$AUGACU*dG*CUC$^{3'}$)] that contains a deoxyguanosine at the site of excision has been shown to prevent the second reaction step of TES (see Chapter 3, Figure 3.4). To test the ability of this new deoxyguanosine substrate to undergo the substrate cleavage reaction, it was allowed to react with the ribozyme under optimized TES reaction conditions (8). We found that the deoxyguanosine substrate (k_{obs} value of 3 ± 0.5 min^{-1}) only minimally reduces the observed rate constant of the substrate cleavage reaction as compared to the normal substrate (k_{obs} value of 3.7 ± 0.2 min^{-1}). Therefore, the deoxyguanosine substrate reasonably mimics the normal substrate as a first reaction step substrate. Importantly, this substrate inhibits the exon-ligation step, allowing us to isolate and analyze only the first reaction step.

Observed Rate Constants for Substrate Cleavage, k_{obs} and k_2

Ribozyme excess conditions were used to determine the pseudo-first-order rate constant for the substrate cleavage reaction. Note that under these reaction conditions the ribozyme-product complex is denatured upon addition of stop buffer, and so product dissociation is not observable. Therefore, these experiments measure the rate of substrate cleavage from the ribozyme-substrate complex.

The observed rate constants (k_{obs}) were measured in reactions containing various ribozyme concentrations (5-300 nM) and 1.3 nM of 5'-end radiolabeled substrate (Figures 4.3A and 4.3B). As seen in Figure 4.3C, the observed rate constants at the higher ribozyme concentrations (100-350 nM) are independent of ribozyme concentration. This indicates that saturation of the ribozyme has been reached. A value of $k_2 = 4.1 \pm 0.5$ min^{-1} and $K_M = 102 \pm 0.4$ nM were obtained by fitting the average k_{obs} values to the Michaelis-Menten equation. Herein, k_2 represents the maximum first order rate of substrate cleavage under single turnover conditions. For lower ribozyme concentrations, 5-40 nM, the k_{obs} values for the substrate cleavage reaction increase linearly with ribozyme concentration. This linear dependence reflects the apparent second order rate constant, k_2/K_M, and the slope gives a value of $(2.8 \pm 0.5) \times 10^7$ M^{-1} min^{-1} (Figure 4.3D). Note the values obtained are similar to values previously reported for other group I intron-derived ribozymes (72, 73, 114).

Dependence of Substrate Cleavage on pH

It has been reported that the rate of substrate cleavage step in *Tetrahymena* (16), *Anabaena* (72), and *Azoarcus* (73) group I introns, as well as reaction steps for some small ribozymes (115-117), show a log-linear increase in the reaction rate constant with increasing pH in the acid range (up to pH 7). This is consistent with a single deprotonation step that activates the cleavage nucleophile, prior to the substrate cleavage reaction (118). This is also consistent with the observed rate constant at a given pH being equivalent to the rate constant of the chemical step at that pH. This was investigated for the *Pneumocystis* ribozyme by measuring the pH dependence of the observed rate constant of the substrate cleavage reaction. As seen in Figure 4.4, the logarithm of the observed rate constant increases linearly with pH in the range of 5 to 7 (slope = 0.5 ± 0.03), but not at higher pH values. In the case of the *Tetrahymena* group I intron-derived ribozyme, such non-linear behavior was attributed to a pH-dependent conformational change occurring within the ribozyme (119, 120). This conformational change thus sets a limit on the observed rate constant of cleavage (k_2), even though the rate constant of chemistry (k_c) is expected to continue to increase with increasing pH (119, 120). Apparently, for our substrate cleavage reaction, the rate of the chemical step is potentially

being masked by a similar conformational change, and so k_2 is not equivalent to k_c. The rate of chemistry (k_c), however, can be approximated by extrapolating the log-linear increase that occurs between pH of 5 and 7 to higher pH values. In our case, k_c is then approximately equal to 5.7 ± 1.1 min^{-1} at pH 7.5 (Figure 4.4).

Control experiments were run to determine whether the observed rate constants shown in Figure 4.4 were being influenced by the specific buffers utilized in the respective reactions. We found that the values obtained using MES and HEPES at pH 6.8 were within standard deviation of each other. This was also true using HEPES and EPPS at pH 7.5. Apparently, there is not a buffer-specific effect on the observed rate constants (k_{obs}). Note that we have not examined the rates of substrate cleavage outside the pH range depicted because protonation or deprotonation of nucleotides is expected to cause general chemical denaturation of the ribozyme (121).

Rate Constant for Substrate Dissociation, k_{-1}

The upper limit of the rate constant for substrate dissociation was measured in a pulse-chase experiment. In this experiment, the time chosen for t_1 (30 s) was such that a significant fraction of the substrate would remain unreacted. After the addition of the chase, which in this case is dilution with buffer, aliquots were removed at designated times up to 15 min. An otherwise identical reaction, but without the added chase, was carried out in parallel. The observed rate constants for the chase reaction ($k_{obs,\ chase}$ = 2.5 ± 0.04 min^{-1}) and in the reaction without added chase ($k_{obs,\ no\text{-}chase}$ = 1.5 ± 0.01 min^{-1}) were obtained from a single-exponential fit of product formation against t_2 (Figure 4.5). The substrate dissociation rate constant (k_{-1} = 0.9 ± 0.04 min^{-1}) was then determined using Equation 2 (see Materials and Methods). Note that k_{-1} is comparable in value to the cleavage step (k_2), implying that the ribozyme-substrate complex does not reach equilibrium with free ribozyme prior to the cleavage step.

Rate Constant for Substrate Association, k_1

The kinetic data indicates that substrate dissociation is comparable in value to the cleavage step. This implies that the second-order rate constant, k_2/K_M, will be a combination of substrate association (k_1), dissociation (k_{-1}) and cleavage (k_2) steps. Thus,

the second-order rate constant can be represented as $k_2/K_M = k_1k_2/(k_{-1} + k_2)$ (122). As discussed earlier, a value of 2.8 x 10^7 M^{-1} min^{-1} was obtained for the second order rate constant k_2/K_M. Using this value of k_2/K_M and the values of k_2 and k_{-1} (4.1 min^{-1} and 0.9 min^{-1} respectively), the calculated value of k_1 is 3.4 x 10^7 M^{-1} min^{-1}.

For confirmation, k_1 was directly measured in a pulse-chase experiment. In this case, ribozyme and radiolabeled substrate were combined for varying times, t_1 (15 s to 120 s), and then chased with excess buffer to prevent further binding. The mixtures were then incubated for a time t_2 = 15 min, which ensures that essentially every substrate molecule that binds to the ribozyme during t_1 is converted to product. The amount of product formed was plotted against time t_1 (Figure 4.6). The k_{obs} values reflect the rate of approach to equilibrium of the ribozyme-substrate complex formation, which is represented by $k_{obs} = k_1$ [E] + k_{-1}. The slope of the plot of k_{obs} versus ribozyme concentration gives the rate of substrate association, k_1 = (1 ± 0.01) x 10^7 M^{-1} min^{-1} (Figure 4.6), which is in reasonable agreement (for ribozyme reactions) with the calculated value above.

Reversibility of the Substrate Cleavage Reaction

Under single turnover conditions, the first order rate constant (k_2) of the substrate cleavage reaction is 4.1 min^{-1} (Figure 4.3), with a typical end point of 70-80%. Over the time period of 15 to 60 min, this end point does not change, indicating that either an internal equilibrium exists between ribozyme-substrate and ribozyme-product complexes or only 70-80% of the substrate is reactive. Such an internal equilibrium has previously been identified in a guanosine-dependent substrate cleavage reaction catalyzed by a *Tetrahymena* group I intron-derived ribozyme (17). Therefore, a pulse-chase experiment was conducted such that this equilibrium, if occurring, could be disturbed and thus detected (17). In this assay, the substrate cleavage reaction was allowed to proceed to completion and then an excess of unlabeled 5' exon mimic was added. The addition of a large excess of unlabeled 5' exon mimic is expected to prevent rebinding of any dissociated radiolabeled substrate or radiolabeled 5' exon reaction product. The result (Figure 4.7) shows that a substantial fraction of the bound radiolabeled product can be converted back to radiolabeled substrate, hence an internal equilibrium exists.

Furthermore, the results imply that product dissociation is slower than or similar to substrate dissociation (17).

Equilibrium Dissociation Constant of the Substrate Cleavage Product, K_d^P, and Substrate, K_d^S

A trace amount of 5'-end labeled substrate cleavage product (the 6-mer) was incubated with various concentrations of ribozyme for 90 min at 44 °C in H10Mg buffer, and the ribozyme-product complex was then partitioned from the unbound product on a native polyacrylamide gel (75, 113). The equilibrium dissociation constant of the 5' exon product (K_d^P = 69 ± 6 nM) was then determined from a plot (Figure 4.8) of the fraction product bound versus ribozyme concentration (75, 113).

The equilibrium dissociation constant of the substrate, K_d^S, is not directly measurable because the substrate cleavage reaction occurs before the binding equilibrium is established. An estimated value of the equilibrium dissociation constant of the substrate can be obtained, however, from the equation $K_d^S = (k_{-1} / k_1) = 90$ nM.

Rate Constant for Dissociation of the 5' Exon Product, k_{-3}

The product dissociation rate constant (k_{-3}) was determined using a pulse-chase assay, combined with native polyacrylamide gel electrophoresis. In this assay, an excess of ribozyme was mixed with 1.3 nM 5'-end labeled 5' exon mimic, which was then incubated in H10Mg buffer containing 3.4 % glycerol at 44 °C for 30 min. An excess amount of unlabeled 5' product was then added to initiate the chase, and aliquots were removed at designated times. These aliquots were directly loaded onto a running native polyacrylamide gel to isolate the bound and unbound fractions. For quantification, the amount of product not bound after the chase was subtracted from that at time t_1, which yields the amount of product dissociated due to the chase. The rate of product dissociation (k_{-3} = 0.09 ± 0.05 min^{-1}) was then obtained from fitting Equation 1 to a single exponential function (Figure 4.9). Apparently, product dissociation is slower than substrate dissociation, which show that the substrate cleavage reaction is not an effective multiple turnover reaction.

DISCUSSION

The *Pneumocystis carinii* ribozyme catalyzes the excision of internal segment from within an RNA substrate (see Figure 4.1). The first reaction step (substrate cleavage) is catalyzed by the 3' terminal guanosine (G336) of the *Pneumocystis* ribozyme (see Chapter 3) and is similar to the cyclization reactions catalyzed by group I introns for the creation of both full-length (FLC) and truncated circular group I introns (67-70, 101). To further enhance our current understanding of the mechanism of the substrate cleavage reaction; a minimal kinetic scheme was developed for the individual steps of the substrate cleavage reaction. The use of a deoxyribose-containing substrate allowed the isolation of the first reaction step (substrate cleavage) while disallowing the second reaction step (exon ligation). Note, however, that in the context of the full TES reaction, the substrate cleavage product is actually an intermediate, between the two reaction steps.

Substrate Binding

The rate constant for the substrate binding the *Pneumocystis* ribozyme, k_1, is far below the diffusional limit of 10^{11} M^{-1} min^{-1} for the collision of small molecules (123). Thus, unlike classical enzymes which react near diffusion-controlled limits (122, 124-126), the *Pneumocystis* ribozyme is not under diffusion control. This value, however, is within the range (10^7 - 10^9 M^{-1} min^{-1}) expected for the formation of RNA duplexes (127-131), as seen with other ribozymes (16, 103, 108, 110, 111, 132). Thus, the rate of assembly of the *Pneumocystis* ribozyme-substrate complex appears to be limited by the process of helix formation. Nevertheless, because k_2/K_M (k_2 = 4.1 min^{-1} and K_M = 102 nM respectively) approaches the rate of substrate association, catalysis can be expected to occur about as fast as base pairing between the ribozyme and substrate. This is typical of ribozymes that bind their substrates through double helices (16, 108, 109, 111, 133). Of equal importance is that the on-rate of the substrate, k_1, is 7 orders of magnitude faster than the off rate, k_{-1}. Therefore, the substrate is relatively slow to dissociate from the ribozyme once bound.

Substrate Cleavage

The rate constant for the substrate cleavage reaction, k_2, under single turnover conditions was determined to be 4.1 min^{-1}. Though it appears that the true rate constant for the actual chemical step is being masked, probably by a local conformational change that occurs after substrate binding and before the actual chemical step, this rate is approximately four times faster than the rate constant for substrate dissociation (k_{-1} = 0.9 min^{-1}). Although the substrate is more likely to react than it is to dissociate, the similar order of magnitude suggests that a non-trivial fraction of the substrate will dissociate before the substrate cleavage reaction occurs.

The 'catalytic power' of an RNA-cleaving ribozyme can be estimated by comparing the observed rate constant of a catalyzed reaction to that of an equivalent uncatalyzed reaction. Under simulated physiological conditions, the uncatalyzed rate constant of the phosphotransesterification reaction (k_{noncat}) is estimated to be 10^{-9} min^{-1} (16, 134). Thus, a rate of 4.1 min^{-1} for the substrate cleavage reaction represents a catalytic rate enhancement (k_2/k_{noncat}) of approximately 10^9-fold. This rate enhancement also corresponds to approximately 13 kcal/mol of transition-state stabilization according to the following equation: $\Delta G° = -RT \ln (k_2/k_{noncat})$, as discussed (16).

It was previously reported that a *Tetrahymena* ribozyme can also catalyze a 3' terminal guanosine-mediated substrate cleavage reaction (13, 15, 17, 68), albeit under different optimized conditions using different substrates. In one such study (17), the 3' terminal guanosine catalyzed reaction was reported to behave similar to the exogenous guanosine catalyzed reaction, for which k_c = 350 min^{-1} (16). In comparison, the *Pneumocystis* endogenous reaction is approximately 60-fold slower (k_c = 5.7 min^{-1}). This substantial difference might be due to the *Tetrahymena* ribozyme being faster than the *Pneumocystis* ribozyme, the difference in reaction conditions, or that the endogenous guanosine nucleophile in the *Pneumocystis* ribozyme, although bound to the GBS, is not bound in an ideal orientation. Indeed, this last idea could be supported in that proper alignment of the endogenous guanosine nucleophile with respect to the *Pneumocystis* ribozyme could be hindered by the absence of a P9.0 helix interaction, which is predicted to align the intramolecular guanosine into the GBS (17). Although more detailed studies using the same reaction conditions and sequences would have to be conducted in order to

make these comparisons more concrete, it is apparent that the intramolecular guanosine-mediated substrate cleavage reaction is substantially slower than either intermolecular or intramolecular guanosine-mediated substrate cleavage reaction for the case of the *Tetrahymena* ribozyme.

Furthermore, the observed rate constant of substrate cleavage shows pH independence between pH 7 and 8.5 implying that at this range the rate of chemistry associated with substrate cleavage is masked by a conformational change (119, 120). The simplest interpretation of this result is that the rate of substrate cleavage is not equivalent to the rate of chemistry, and that the rate of chemistry (extrapolated to be $k_c = 5.7\ min^{-1}$) is faster than the rate of substrate cleavage (measured to be $k_2 = 4.1\ min^{-1}$).

Product Dissociation

For the fraction of substrates that do undergo the reaction, the resultant products dissociate from the ribozyme relatively slowly on the time scale of the reaction. Furthermore, dissociation of the 5' exon product is slower than the cleavage step (by about 45-fold), which significantly impedes the ribozyme from catalyzing multiple turnover reactions. Of course, the 5' exon product of the cleavage reaction is an intermediate in the complete TES reaction, and so slow product dissociation is beneficial for the TES reaction as a whole. In addition, the product off-rate, k_{-3}, is 20-fold slower than the substrate off-rate, k_{-1}. Apparently, there are additional or more stable interactions that the ribozyme uses to bind the product relative to the ribozyme binding the substrate. This is perhaps due to destabilization of substrate binding via positioning of the 3' bridging phosphoryl oxygen at the cleavage site next to a required Mg ion in the ground state (135). As negative charge develops on the 3' oxygen upon entering the transition state, this interaction will become more favorable. This transition state stabilization is thought to be an important stabilization/destabilization factor in ribozyme-substrate binding (135).

A Conformation Change Exists Between the Two Steps of the TES Reaction

The substrate guanosine to be excised (G_1) and its 2'-OH group are required for the second step of TES (18), similar to (if not the same as) the role of the ωG in the

second step of self-splicing (57, 63, 136-141). This suggests that the G_1 is likely binding to the GBS of the ribozyme prior to the exon-ligation step of TES. In the substrate cleavage step, however, the 3' terminal guanosine (G336) of the ribozyme (Figure 4.1) is binding to that same GBS. Therefore, for the TES reaction, there is likely a local conformational change between the two reaction steps that sees the G_1 displace the ribozymes 3' terminal guanosine for binding into the GBS (Figure 4.1). That there exists a conformational change is further supported by kinetic data that shows that for the substrate cleavage reaction, substrate dissociation is faster than product dissociation (Figure 4.2), implying that different (strengths of) interactions hold the substrate and the 5' exon cleavage product. The local conformational change that occurs in TES is likely similar to the local conformational change that occurs in self-splicing, with the displacement of the exogenous guanosine by the ωG of the intron (3, 46, 47, 59). Nevertheless, as TES uses an intramolecular nucleophile and self-splicing uses an intermolecular nucleophile, the local conformational changes between the two steps of each reaction are likely not identical.

Implications for TES applications

TES substrates, once bound, are four times more likely to undergo the substrate cleavage reaction as they are to dissociate. Therefore, to make more effective TES ribozymes, one could decrease the rate of substrate dissociation relative to that for the substrate cleavage reaction. Potential strategies for achieving this are to increase the strength of helix P1, either through target selection or elongation of helix P1. Note, however, that this strategy could result in a decrease in the rate of substrate cleavage.

The results also suggest that the *Pneumocystis* ribozyme catalyzes the substrate cleavage reaction (catalyzed by either an exogenous or endogenous guanosine) approximately 60-fold slower than the *Tetrahymena* ribozyme. These results suggest that the rate of the substrate cleavage reaction could be improved upon by decreasing the rate of substrate dissociation before the cleavage step.

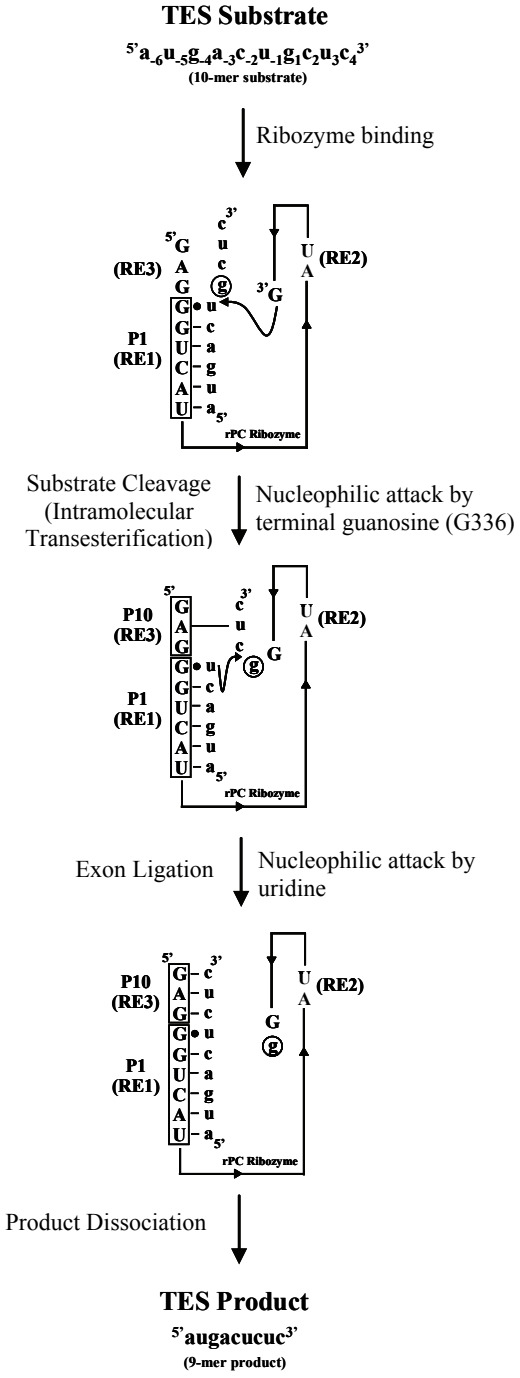


Figure 4.1. Schematic of the two-step TES reaction. The rPC ribozyme is in uppercase lettering and, the 10-mer substrate is in lowercase lettering, and the ribozyme recognition elements RE1 and RE3 base pair with the substrate to form helices P1 and P10, respectively. The sites of catalysis for the first step (substrate cleavage) and the second

step (exon ligation) are shown with arrows, and the guanosine to be excised (G_1) is circled. The diagram shows only the recognition elements of the ribozyme.

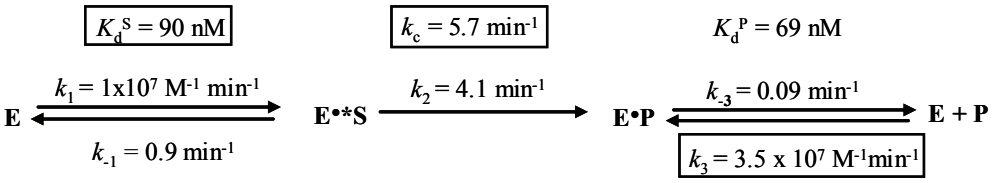


Figure 4.2. Kinetic scheme for the substrate cleavage reaction. E denotes the rPC ribozyme, S denotes the 10-mer substrate, and P denotes the 6-mer cleavage product. All rate and equilibrium constant values were measured or calculated (boxed values) in this report. The scheme includes rate constants for substrate association (k_1) and dissociation (k_{-1}), cleavage (k_2), and product association (k_3) and dissociation (k_{-3}). Note that the observed rate constant for the cleavage step (k_2) is distinguishable from the actual rate constant for chemistry (k_c). The scheme also includes equilibrium constants for substrate (K_d^S) and product (K_d^P) dissociation.

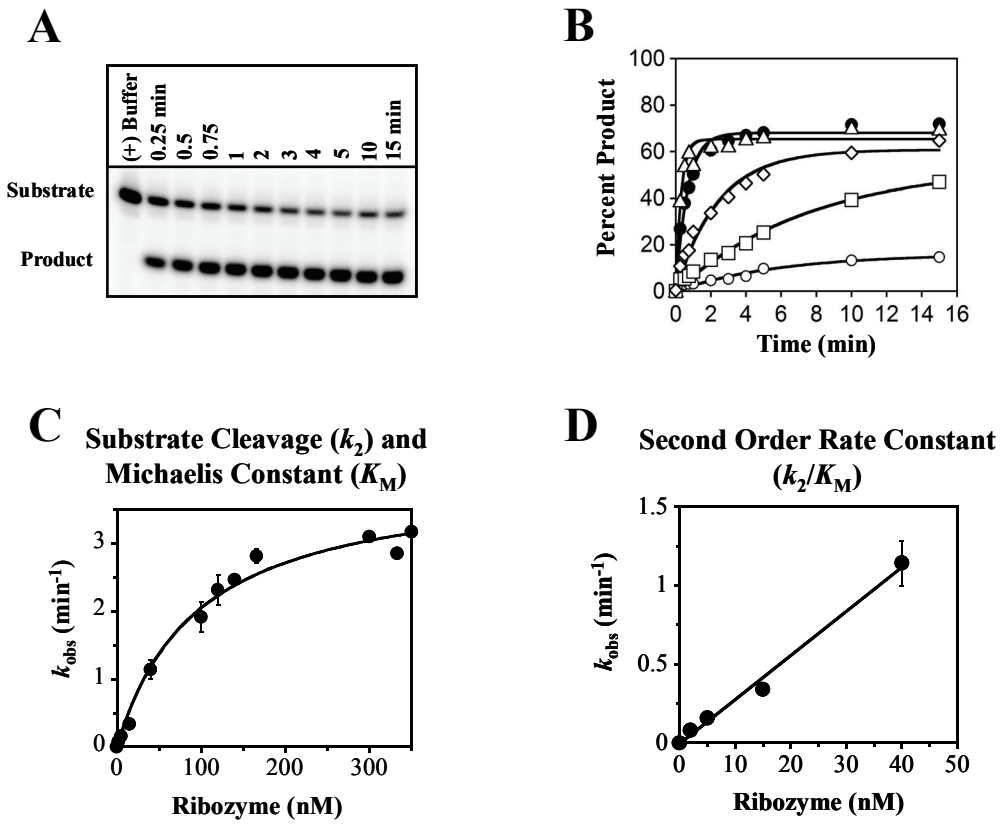


Figure 4.3. Substrate cleavage reactions. All reactions were conducted in H10Mg buffer. (A) Representative polyacrylamide gel with the 5'-end labeled substrate and 166 nM rPC ribozyme. The positions of the substrate and the substrate cleavage product on the gel are labeled. The lane marked (+) buffer contains a 15 min reaction in the absence of the ribozyme. (B) Representative plot of the substrate cleavage reaction at 5 nM (▲), 10 nM (○), 20 nM (□), 40 nM (◊), 166 nM (Δ), and 300 nM (●) ribozyme concentration. All data points between the two independent assays have a standard deviation less than 15%. (C) Fit of the average k_{obs} values from panel B versus ribozyme concentration (0-350 nM) to the Michaelis-Menten equation. The plot resulted in a value of k_2 = 4.1 ± 0.5 min^{-1} and K_M = 102 ± 0.4 nM respectively. These values are the average of the two independent assays. D. Linear fit of the average k_{obs} values from panel B versus ribozyme concentration (5-40 nM). The resulting k_2/K_M value (2.8 ± 0.5 x 10^7 M^{-1} min^{-1}) is the average of the two independent assays.

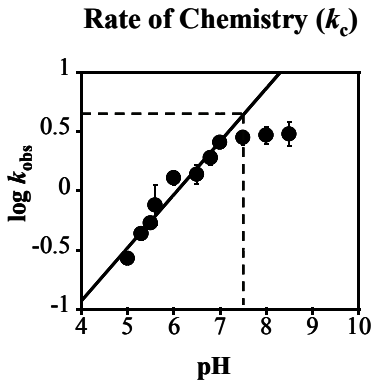


Figure 4.4. pH dependence of the observed rate of substrate cleavage. Values for were obtained from single turnover reactions at 44 °C using 166 nM ribozyme, 1.3 nM 5'-end labeled substrate in buffer containing 135 mM KCl and 10 mM $MgCl_2$. The buffers used for this study were 50 mM MES (pH 5.0-6.8), 50 mM HEPES (pH 7.0, 7.5), or 50 mM EPPS (pH 8.0, 8.5). Each rate is the average of two independent measurements and has a standard deviation less than 15%. The solid line represents the log-linear increase in the data set from pH of 5 to 7 (slope = 0.5 ± 0.03). The extrapolation of the line to pH 7.5 (depicted by dashed lines) gives a value of 0.75 ± 0.1 which corresponds to rate of chemistry (k_c) of 5.7 ± 1.1 min^{-1}.

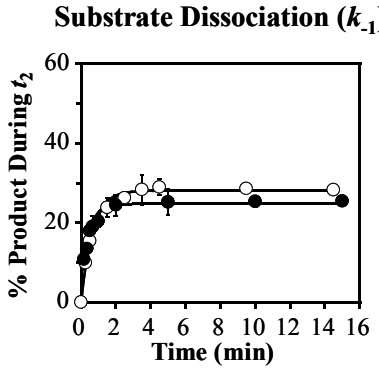


Figure 4.5. Determination of the rate constant for substrate dissociation, k_{-1}. (A) Scheme of the pulse-chase experiment, which was conducted in H10Mg buffer at 44 °C and 166 nM ribozyme. The chase was initiated by diluting the reaction mixture with H10Mg buffer. (B) Representative plot of cleaved substrate, after t_1, versus time (t_2) with chase (closed circles) and without added chase (open circles). The resultant first order rate constants obtained with the chase ($k_{obs,\ chase}$ = 2.5 ± 0.04) and without the chase ($k_{obs,\ no\text{-}chase}$ = 1.5 ± 0.01 min^{-1}) are the average of two independent assays. All data points between the two independent assays have a standard deviation less than 10%. From this data, the rate of substrate dissociation, k_{-1}, is 0.9 ± 0.04 min^{-1}.

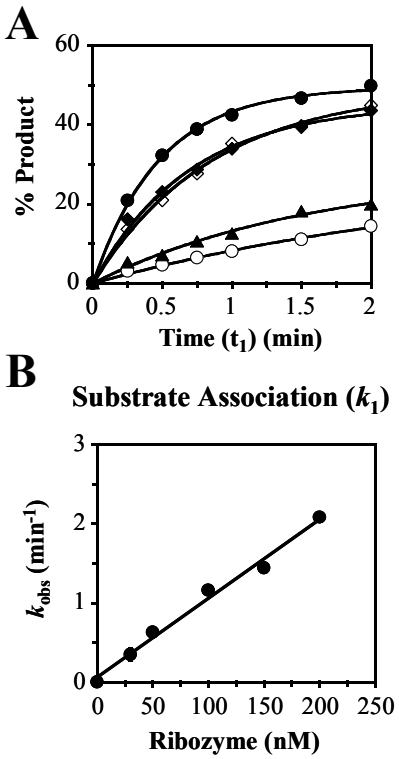


Figure 4.6. Determination of the rate constant for substrate association, k_1. (A) Representative plot of pulse-chase experiments in H10Mg buffer at 44 °C with five different ribozyme concentrations: 30 nM (○), 50 nM (▲), 100 nM (◊), 150 nM (♦), and 200 nM (●). All data points between the two independent assays have a standard deviation less than 10%. (B) Representative plot of the k_{obs} values against ribozyme concentration. The line is fit to the equation $k_{obs} = k_1[E] + k_{-1}$ and the substrate association rate ($k_1 = 1 \pm 0.01 \times 10^7$ M^{-1} min^{-1}) was calculated from the slope. Note that the error bars present on the graph are too small to be seen.

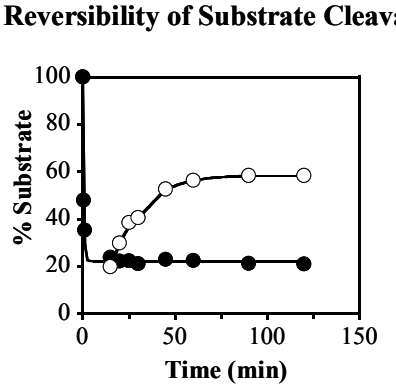


Figure 4.7. Substrate cleavage products undergo the reverse reaction. A plot of the disappearance of substrate in a normal reaction (no chase, closed circles) and reappearance of the substrate in the presence of chase (open circles). Each point on the plot is the average of two independent experiments, and has a standard deviation less than 10%. Note that the error bars present on the graph are too small to be seen.

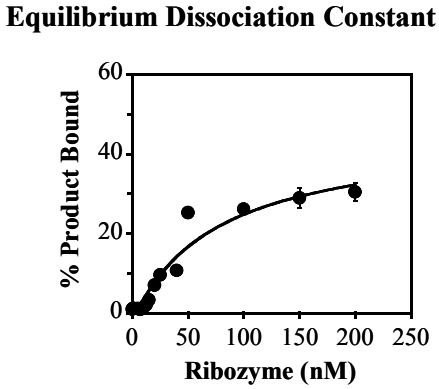


Figure 4.8. Determination of the equilibrium dissociation constant of the substrate cleavage product, K_d^P. In the reaction, various concentrations of ribozyme were mixed with trace amounts of 5'-end radiolabeled substrate cleavage product in H10Mg buffer containing 3.4 % glycerol. Shown is a representative plot of the percent substrate cleavage product bound to the ribozyme versus ribozyme concentration. The resultant value of K_d^P is 69 ± 6 nM is the average of two independent assays with each data point having a standard deviation less than 15%.

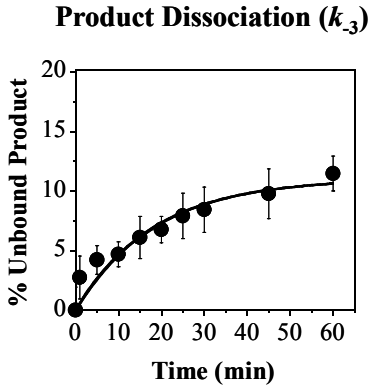


Figure 4.9. Determination of the rate constant for dissociation of the 5' exon product, k_{-3}. Representative plot of the fraction of unbound product versus chase time, t_2. The rate of product dissociation, k_{-3}, is 0.09 ± 0.05 min^{-1}, which is the average of two independent assays with each data point having a standard deviation typically less than 20%.

CHAPTER FIVE- THE TRANS EXCISION-SPLICING REACTION CAN BE CATALYZED BY GROUP I INTRON-DERIVED RIBOZYMES FROM *TETRAHYMENA THERMOPHILA* AND *CANDIDA ALBICANS*

INTRODUCTION

Group I intron-derived ribozymes, which are essentially exon-less introns, have previously been shown to catalyze non-native reactions that are similar to the self-splicing reaction (3, 6). In one such reaction, termed trans excision-splicing (TES), an intron-derived ribozyme from *Pneumocystis carinii* catalyzes the excision of an internal segment from within an RNA substrate, resulting in the reattachment of the flanking regions (Figure 5.1) (8). The TES reaction, which is similar to other ribozyme-mediated reactions, proceeds through two consecutive transesterification reactions: substrate cleavage followed by exon ligation (Figure 1). For the substrate cleavage reaction, the 3' terminal guanosine of the *Pneumocystis* ribozyme catalyzes the intramolecular cleavage of TES substrates (Figure 5.1) (see Chapter 3). Substrate cleavage is followed by exon ligation, resulting in formation of the final TES products (Figure 5.1).

In general, group I intron-derived ribozymes contain all the structural elements required for catalyzing the TES reaction. This suggests that group I intron-derived ribozymes other than *P. carinii* could catalyze the TES reaction. Therefore, intron-derived ribozymes from *Tetrahymena thermophila* and *Candida albicans* were constructed and tested for their ability to catalyze TES reactions, using as substrates their own native 5' and 3' exon sequences (Figure 5.2) (71, 74).

In this study, the *Tetrahymena* and *Candida* ribozymes were shown to catalyze the excision of a single nucleotide from within their ribozyme-specific substrates. Furthermore, both ribozymes were shown to catalyze substrate cleavage through intramolecular transesterification, similar to the *Pneumocystis* ribozyme. The overall yield of TES reactions, however, differed between the individual ribozyme constructs. The *Tetrahymena* and *Candida* ribozyme constructs were shown to catalyze the individual steps of TES less efficiently (~20% lower yields) than the *Pneumocystis* ribozyme. This variation appears to be due to the individual ribozymes ability to

productively utilize their 3' terminal guanosine as the intramolecular nucleophile. Overall, however, the results demonstrate TES is a general ribozyme-mediated reaction.

MATERIALS AND METHODS

Oligonucleotide Synthesis and Preparation

DNA oligonucleotides were obtained from Integrated DNA Technologies (Coralville, IA) and were used without further purification. RNA oligonucleotides were obtained from Dharmacon Research, Inc. (Lafayette, CO) and were deprotected using the manufacturer's suggested protocol. Concentrations of individual oligonucleotides were calculated from UV absorption measurements using a Beckman DU 650 spectrophotometer (Beckman Coulter, Inc.; Fullerton, CA). RNA oligonucleotides were 5'-end radiolabeled by phosphorylation of the 5' terminal hydroxyl group with [γ-^{32}P] ATP (Amersham Pharmacia Biotech; Piscataway, NJ) using T4 polynucleotide kinase (New England Biolabs; Beverly, MA) as previously described (75). RNA oligonucleotides were 3'-end radiolabeled by ligation of radiolabeled pCp (5'-*pCp) to the 3'-end of the RNA oligonucleotide using T4 RNA ligase (New England Biolabs), as previously described (10). Note that 3'-end radiolabeling results in substrates being one nucleotide longer than 5'-end radiolabeled substrates.

Plasmid Construction

For these studies, two ribozyme constructs were created: pT-X, derived from the plasmid pT7L-21 (kindly provided by Douglas H. Turner, University of Rochester; Rochester, NY) which contains the *Tetrahymena* ribozyme (71), and pC, derived from the plasmid pC10/1x (provided by Douglas H. Turner) that contains the *Candida* ribozyme (74). The individual ribozyme sequences were modified to include an RE3 sequence so that a P10 helix interaction would form upon binding substrate (see Figure 5.2).

The plasmids, pT and pC, were made by changing the parental plasmids, pT7L-21 and pC10/1x, via site directed mutagenesis. The pT plasmid was created using a single round of site directed mutagenesis using the primer pairs (underlined bases represent

inserted nucleotides) $^{5'}$CGACTCACTATAGACCTTTGGAGGGAAAAG$^{3'}$ and $^{5'}$CTTTTCCCTCCAAAGGTCTATAGTGAGTCG$^{3'}$ for RE3 formation. The pC plasmid was created using a single round of site-directed mutagenesis using the primer pairs $^{5'}$CGACTCACTATAGGTAAGGGAGGCAAAAGTAGGG$^{3'}$ and $^{5'}$CCCTACTTTTGCCTCCCTTACCTATAGTGAGTCG$^{3'}$ for RE3 formation. Note that a guanosine was inserted at the 5'-end of the each ribozyme construct to ensure efficient transcription by T7 RNA polymerase (92).

Additionally, an *Xba*I restriction site was inserted at the 3'-end of the pT plasmid, which creates a terminal guanosine (required for the intramolecular transesterification reaction), as well as making the 3'-ends of all the ribozymes the same (see Chapter 3). The plasmid, pT-X, was created in two rounds using the following primer pairs (underlined bases represent changes to the 3'-end of the ribozyme) $^{5'}$GGAGTACTCGATAAGGTAGCCAAATGCC$^{3'}$ and $^{5'}$GGCATTTGGCTACCTTATCGAGTACTCC$^{3'}$ for the first round. The primer pairs $^{5'}$GGAGTACTCTAGATAAGGTAGCCAAATGCC$^{3'}$ and $^{5'}$GGCATTTGGCTACCTTATCTAGAGTACTCC$^{3'}$ were used for the second round.

Each set of primers (22.5 pmol for each primer) was used in an amplification reaction consisting of 25 ng parental plasmid (pT7L-21, pT, or pC10/1x), 2.5 units Pfu DNA polymerase (Stratagene; La Jolla, CA), and 0.5 mM dNTPs (Invitrogen; Grand Island, NY) in a buffer consisting of 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 20 mM Tris-HCl (pH 8.8), 2 mM $MgSO_4$, 0.1% Triton X-100, and 0.1 mg/mL BSA (total reaction volume is 50 μL). After an initial denaturation for 30 s at 95 °C, the mixture was subjected to 20 cycles of 95 °C for 30 s, 50 °C for 2 min, and 68 °C for 6 min. Parental plasmid was then digested with 20 units *Dpn*I (New England Biolabs) in 4.1 μL of the manufacturer's supplied buffer (10X NEBuffer 4) for 2 h at 37 °C. A 3 μL aliquot of this mixture was then used to transform *Escherichia coli* DH5α competent cells (Gibco BRL). The vectors were purified using a QIAprep Spin Miniprep Kit (QIAGEN; Valencia, CA), and were sequenced for confirmation of the individual changes (Davis Sequencing; Davis, California).

Ribozyme Synthesis and Purification

The *Pneumocystis* plasmid (PC) (8), the *Tetrahymena* plasmid (pT-X), and the *Candida* plasmid (pC) were linearized in a 50 µL reaction consisting of 16 µg plasmid, 5 µL 10X REACT 2, and 50 U of *Xba*I (Invitrogen; Grand Island, NY). The reactions were incubated at 37 °C for 2 h followed by purification using a QIAquick PCR purification kit (Qiagen) using the manufacturers suggested protocol. The resulting linearized plasmids were subjected to run-off transcription reactions in a 50 µL volume consisting of 1 µg DNA template, 50 U of T7 RNA polymerase (New England Biolabs), 40 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 5 mM DTT, 5 mM spermidine, and 1 mM rNTP mix. The reactions were incubated at 37 °C for 2 h, at which time each ribozyme construct was purified using a Qiagen Plasmid Midi Kit (Qiagen), as previously described (8). Concentrations of individual ribozymes were calculated from UV absorption measurements using a Beckman DU 650 spectrophotometer (Beckman Coulter, Inc.) using the previously reported extinction coefficient of 3.5 x 10^6 M^{-1} cm^{-1} for the *Pneumocystis* (rPC) ribozyme (75), 3.2 x 10^6 M^{-1} cm^{-1} for the *Tetrahymena* (rT-X) ribozyme (71), and 3.72 x 10^6 M^{-1} cm^{-1} for the *Candida* (rC) ribozyme (74).

TES Reactions

Reactions were performed at 44 °C in buffer consisting of 50 mM Hepes (25 mM Na^+), 135 mM KCl, and various concentrations of $MgCl_2$. TES reactions were optimized for magnesium concentration (0-35 mM $MgCl_2$), ribozyme concentration (0-500 nM), and time (1-45 min). In each case, TES reaction conditions were optimized to give the greatest TES product yield. Briefly, a ribozyme solution, in HXMg buffer, lacking substrate was preannealed at 60 °C for 5 min, and then slow cooled to 44 °C. Reactions were initiated by the addition of 1 µL of 8 nM 5'-end radiolabeled substrate (10-mer, $^{5'}$AUGACU*G*CUC$^{3'}$ for rPC; 13-mer, $^{5'}$CUCUCU*G*AAAGGU$^{3'}$ for rT-X; or 11-mer, $^{5'}$GACUCU*G*CUUA$^{3'}$ for rC). Note that the italic lettering in the substrates represent the nucleotide to be excised. The final concentration of the substrates in the reaction mixtures was 1.3 nM. For mechanistic studies, 3'-end radiolabeled substrates were employed at the same concentrations stated above. Reactions were conducted at 44 °C for the optimized amount of time, and stopped by adding an equal volume of stop buffer (10 M urea, 3 mM

EDTA, and 0.1X TBE). Reactions were denatured for 1 min at 90 °C and substrates, intermediates, and products were separated on a 12% polyacrylamide/8 M urea gel. The gel was transferred to chromatography paper (Whatman 3MM CHR) and dried under vacuum. The bands were visualized and quantified on a Molecular Dynamics Storm 860 Phosphorimager.

Kinetic Studies

The observed rate constant (k_{obs}) for the individual steps of the TES reaction, using each ribozyme construct, was obtained under single-turnover 'ribozyme excess' conditions. In these experiments, 35 μL of ribozyme in appropriate buffer was preannealed at 60 °C for 5 min and slow cooled to 44 °C. The reactions were initiated by the addition of 7 μL of 8 nM 5'-end radiolabeled substrate (10-mer, $^{5'}$AUGACU*G*CUC$^{3'}$ for rPC; $^{5'}$CUCUCU*G*AAAGGU$^{3'}$ for rT-X; or $^{5'}$GACUCU*G*CUUA$^{3'}$ for rC), bringing the final concentration to 1.3 nM substrate. A 3 μL aliquot was periodically removed and added to an equal volume of stop buffer. The reactants and products were separated on a 12% polyacrylamide/8 M urea gel, and the gel was dried under vacuum. The resulting bands were visualized and quantified as described above. The observed rate constants for the substrate cleavage step (substrate cleavage plus TES product over time) and for exon ligation (TES product over time) were quantified as previously described (8).

RESULTS

To simplify the comparison of the individual ribozyme constructs, our simplest TES model test system was utilized. Successful TES reactions will result in the excision of a single nucleotide from within their ribozyme-specific substrates. For these studies, the previously constructed *Tetrahymena* and *Candida* ribozymes were modified in two ways (see Figure 5.2). First, the initial ribozyme constructs lacked the nucleotides required to form the P10 helix (RE3), and so an RE3 region (native to the ribozyme) was added. As seen for the *Pneumocystis* ribozyme, P10 helix formation has been shown to aid in binding the 3' exon region of TES substrates (18). Second, the 3'-end of each

ribozyme sequence was modified to resemble the *Pneumocystis* ribozyme (see Chapter 3), including the addition of a terminal guanosine as it is required as the intramolecular nucleophile in the substrate cleavage reaction (Chapter 3).

TES Reactions Conducted with the Pneumocystis, Tetrahymena, and Candida Ribozymes.

As a control, the TES reaction using the previously characterized *Pneumocystis* (rPC) ribozyme (8) was run (Figure 5.2). TES reactions conducted with 5'-end radiolabeled substrate ($^{5'}$AUGACU*G*CUC$^{3'}$, underlined nucleotides are excised in TES reactions) resulted in the expected 9-mer product ($^{5'}$AUGACUCUC$^{3'}$) and 6-mer intermediate ($^{5'}$AUGACU$^{3'}$) (Figure 5.3A, Lanes G-H). TES reaction conditions were optimized for magnesium concentration, time, and ribozyme concentration for maximum TES product yields (Figure 5.3B). Optimized TES reaction conditions resulted in 68.5 ± 4.4% TES product yield, with an observed rate constant, k_{obs}, of 5.0 ± 1.1 min^{-1} for substrate cleavage and 3.2 ± 0.1 min^{-1} for exon ligation. Note, optimized TES reactions conditions are compiled in Table 5.1.

For the *Tetrahymena* (rT-X) ribozyme, TES reactions conducted with 5'-end radiolabeled substrate ($^{5'}$CUCUCU*G*AAAGGU$^{3'}$, underlined nucleotides are excised in TES reactions) resulted in the predicted 12-mer product ($^{5'}$CUCUCUAAAGGU$^{3'}$) and 6-mer intermediate ($^{5'}$CUCUCU$^{3'}$) (Figure 5.4A, Lanes G-H). Optimized TES reactions resulted in 47.5 ± 0.2% TES product yield (Figure 5.4B), with an observed rate constant, k_{obs}, of 0.17 ± 0.05 min^{-1} for substrate cleavage and 0.14 ± 0.02 min^{-1} for the exon ligation (Table 5.1).

For the *Candida* (rC) ribozyme, TES reactions conducted with 5'-end radiolabeled substrate ($^{5'}$GACUCU*G*CUUA$^{3'}$, underlined nucleotides are excised in TES reactions) resulted in the predicted 10-mer product ($^{5'}$GACUCUCUUA$^{3'}$) and 6-mer intermediate ($^{5'}$GACUCU$^{3'}$) (Figure 5.5A, Lanes G-H). Optimized TES reactions conditions resulted in 50 ± 1.3% TES product yield (Figure 5.5B), with an observed rate constant, k_{obs}, of 0.31 ± 0.01 min^{-1} for substrate cleavage reaction and 0.34 ± 0.03 min^{-1} for exon ligation (Table 5.1). These results demonstrate that the *Tetrahymena* and *Candida* ribozymes can catalyze the excision of a single nucleotide from within their ribozyme-specific substrates, through the TES reaction.

The Tetrahymena and Candida Ribozymes Catalyze the Substrate Cleavage Reaction through Intramolecular Transesterification.

In the absence of an exogenous guanosine cofactor, the substrate cleavage reaction can be catalyzed through two distinct reaction mechanisms. These include ribozyme-mediated 5' hydrolysis (12, 16) and intramolecular transesterification (13, 15, 17, 68). For the *Pneumocystis* ribozyme, the 3' terminal guanosine catalyzes substrate cleavage through an intramolecular transesterification reaction (see Chapter 3). Given that the *Tetrahymena* and *Candida* ribozyme constructs were synthesized with 3' terminal guanosines, it is expected that each of these constructs will catalyze their substrate cleavage reaction through intramolecular transesterification. With these modifications, the two ribozymes would be predicted to catalyze substrate cleavage through intramolecular transesterification.

As seen for the *Pneumocystis* ribozyme, the two substrate cleavage reaction mechanisms can be distinguished by utilizing 3'-end radiolabeled substrates. This is because the size of the TES reaction intermediates will differ depending upon whether the substrate cleavage step is hydrolysis or intramolecular transesterification (Chapter 3). For the *Tetrahymena* ribozyme, TES reactions were conducted with 3'-end labeled substrate ($^{5'}$CUCUCU*G*AAAGGU<u>C</u>$^{3'}$), which resulted in the predicted 13-mer product ($^{5'}$CUCUCUAAAGGU<u>C</u>$^{3'}$), as well as the predicted higher molecular weight product band (Figure 5.4A, Lanes I-J). The absence of the predicted hydrolysis product ($^{5'}$GAAAGGUC$^{3'}$), suggests the *Tetrahymena* ribozyme undergoes intramolecular transesterification.

For the *Candida* ribozyme, TES reactions were conducted using 3'-end radiolabeled substrate ($^{5'}$GACUCU*G*CUUA<u>C</u>$^{3'}$), which resulted in the predicted 11-mer product ($^{5'}$GACUCUCUUA<u>C</u>$^{3'}$), as well as the predicted high molecular weight product band (Figure 5.5A, Lanes I-J). Again the absence of the predicted hydrolysis product (6-mer, $^{5'}$GCUUAC$^{3'}$), suggests that the *Candida* ribozyme catalyzes substrate cleavage through intramolecular transesterification.

As a control, TES reactions were conducted with the *Pneumocystis* ribozyme using 3'-end labeled substrate ($^{5'}$AUGACU*G*CUC<u>C</u>$^{3'}$). TES reactions resulted in the

expected 12-mer product ($^{5'}$AUGACUCUCC$^{3'}$) and the expected high molecular weight product band (Figure 5.3A, lanes I-J).

Competition Between Exogenous Guanosine and Endogenous Guanosine for the GBS Inhibits the TES Reaction.

Apparently, the *Tetrahymena* and *Candida* ribozymes catalyze the substrate cleavage reaction through an intramolecular transesterification reaction mechanism. The *Pneumocystis* ribozyme catalyzes the substrate cleavage reaction ~20-fold faster, however, than either the *Tetrahymena* or *Candida* ribozyme (see Table 5.1). This observed decrease in rate could be directly related to the ribozymes' ability to utilize their 3' terminal guanosines for substrate cleavage. For a previously described *Tetrahymena* ribozyme (17), competition assays using exogenous guanosine, were used as a probe for the ribozymes ability to productively bind their terminal guanosine. In these assays, the exogenous guanosine can compete with the ribozymes terminal guanosine for binding to the GBS. Using a similar strategy, inhibition of TES product formation could be directly correlated to the ribozymes ability to effectively bind its terminal guanosine into the GBS.

For the *Pneumocystis* ribozyme, TES reactions were conducted with increasing concentrations of rGTP under optimal TES reaction conditions (Table 5.1). The results demonstrate that TES product formation is not completely hindered until excessively high concentrations of rGTP (10 mM) are added (Figure 5.3C). For the *Tetrahymena* ribozyme, TES reactions resulted in complete inhibition of TES product formation at much lower rGTP concentrations (0.02 mM) (Figure 5.4C). In the case of the *Candida* ribozyme, TES reactions demonstrate that low concentrations of rGTP (0.04 mM) are also effective at completely hindering TES product formation (Figure 5.5C). These results suggest that the *Pneumocystis* ribozyme binds its terminal guanosine (G336) more effectively than either the predicted terminal guanosine of the *Tetrahymena* ribozyme (G416) or the *Candida* ribozyme (G374).

DISCUSSION

In this report, group I intron-derived ribozymes from *Tetrahymena* and *Candida* were shown to be capable of catalyzing the excision of internal segments (in this case a single nucleotide) from within their native 5' and 3' exon sequences. Furthermore, the *Tetrahymena* and *Candida* ribozyme constructs were shown to catalyze the substrate cleavage reaction through an intramolecular transesterification mechanism, using their predicted 3' terminal guanosines as the nucleophiles. This is the first report that group I intron-derived ribozymes other than *Pneumocystis* can catalyze the excision of internal segments from RNA transcripts via the TES reaction.

The Tetrahymena Ribozyme

The *Tetrahymena* and *Pneumocystis* ribozymes catalyze the individual steps of the TES reaction under similar optimized magnesium, time, and ribozyme conditions (see Table 5.1). The results, however, show that the *Pneumocystis* ribozyme catalyzes the individual steps of the TES reaction with higher yields than the *Tetrahymena* ribozyme. Furthermore, the *Pneumocystis* ribozyme catalyzes substrate cleavage ~30-fold faster than the *Tetrahymena* ribozyme (Table 5.1). In addition, the *Pneumocystis* ribozyme catalyzes exon ligation ~20-fold faster than the *Tetrahymena* ribozyme.

One reason that could explain the differences seen for the *Tetrahymena* and *Pneumocystis* ribozymes is that additional nucleotides, other than guanosine, are being added to the 3'-end of these ribozymes during *in vitro* transcription. The non-encoded addition of nucleotides to the 3'-end of RNA transcripts has previously been shown to be a consequence of run-off transcription utilizing T7 RNA polymerase (92). The *Pneumocystis* ribozyme, however, has been shown to catalyze its own 3'-end activation, which would result in the removal of these additional nucleotides (see Chapter 3). In contrast, previously characterized *Tetrahymena* ribozymes that catalyze intramolecular transesterification reactions had to be activated by subjecting the ribozyme to a site-specific 3' hydrolysis reaction (13, 15, 17, 97). It is unknown whether this pre-activation, which is not necessary for *Pneumocystis*, would have aided in the *Tetrahymena* reaction. Therefore, the *Tetrahymena* ribozyme may be hindered in its ability to catalyze site-

specific 3' hydrolysis reactions under the prevailing TES reaction conditions, which could lead to a lower population of active ribozyme.

Alternatively, or in addition, the *Tetrahymena* ribozyme may catalyze reactions different from TES. For instance, a higher molecular weight product band is present with TES reactions conducted with 5'-end labeled substrates (Figure 5.4A). This high molecular weight product band is predicted to be the result of a previously characterized ribozyme-mediated reverse-splicing reaction (79, 80). In this reaction, the ribozyme becomes embedded between the 5' and 3' exon regions of its RNA substrate. Note that this higher molecular weight product band is absent for TES reactions conducted with the *Pneumocystis* ribozyme. That the *Tetrahymena* ribozyme can potentially catalyze this reaction is due to the addition of a guanosine to the 5'-end of the ribozyme coding sequence (added to facilitate efficient *in vitro* transcription). If in fact the reverse-splicing is occurring, then two competing reactions (TES and reverse-splicing) will simultaneously be occurring, which could explain the lower efficiency of the *Tetrahymena* ribozyme reactions.

In addition, previously characterized *Tetrahymena* ribozymes (68, 69, 142), which closely resemble our *Tetrahymena* ribozyme construct, are susceptible to both hydrolysis and intramolecular cyclization reactions in the absence of substrate (68, 69, 142). Both of these reactions result in either a truncated or circular ribozyme product (68, 69, 142). If either of these reactions is occurring, the result would be a reduction of fully functional *Tetrahymena* ribozyme. Although, we have no evidence for these alternative reactions, they could be occurring in small amounts, which would partially explain the lower efficiency for the *Tetrahymena* ribozyme reactions.

The Candida Ribozyme

The *Candida* and *Pneumocystis* ribozymes catalyze the individual steps of the TES reaction under similar optimized time and ribozyme conditions (Table 5.1). For the *Candida* ribozyme, however, higher magnesium concentrations are required for optimal TES reactivity. Similar to the *Tetrahymena* ribozyme, the *Candida* ribozyme was observed to catalyze the individual steps of the TES reaction with lower yields than the *Pneumocystis* ribozyme. Moreover, the *Candida* ribozyme catalyzes substrate cleavage

~15-fold slower than the *Pneumocystis* ribozyme (Table 5.1). In addition, the *Candida* ribozyme catalyzes exon ligation ~10-fold slower than the *Candida* ribozyme.

As for the *Tetrahymena* ribozyme, it is not known if a substantial percentage of the population of *Candida* ribozymes terminates in guanosine. Similarly, it is not known if the *Candida* ribozyme can catalyze a site-specific 3' hydrolysis reaction prior to becoming activated. If the *Candida* ribozyme is not being activated under prevailing TES reaction conditions, this could explain the lower efficiency of the *Candida* ribozyme reactions.

Lastly, the *Candida* ribozyme produces a high molecular weight product band, similar to *Tetrahymena* ribozyme, for TES reactions conducted with 5'-end labeled substrates (Figure 5.5A). At higher magnesium concentrations, the amount of this high molecular weight product band increases with increasing magnesium concentrations. As stated above, if in fact TES and reverse-splicing are simultaneously occurring then the lower efficiency observed for the *Candida* ribozyme would be predicted.

Productive Docking of the 3' Terminal Guanosine Effects the Overall TES Reaction.

Prior to the substrate cleavage reaction, it is predicted that the ribozymes 3' terminal guanosine will be bound into the GBS (17, 68). From the competition studies, it appears that the terminal guanosine of the *Pneumocystis* ribozyme is more productively bound to the GBS than the *Tetrahymena* or *Candida* ribozymes. These results suggest that the decrease in binding affinity of the *Tetrahymena* and *Candida* ribozymes is possibly due to their inability to bind the terminal guanosine into the GBS. This is surprising in that the 3'-end of the ribozyme constructs are all similar.

As seen for ribozyme-mediated intramolecular transesterification reactions, the 3' terminal guanosine is predicted to be aligned into the GBS through helix P9.0 (17, 68). For example, a previously described *Tetrahymena* ribozyme, L21-G414, was predicted to align its 3' terminal guanosine into the GBS through the aid of P9.0 helix formation (17). For L21-G414 ribozyme construct, the sequence at the 3'-end of the ribozyme was identical to the native intron sequence, thereby allowing the two nucleotides preceding the terminal guanosine to be involved in P9.0 helix formation. For our studies, however, the 3' terminal sequence of the *Tetrahymena* and *Candida* ribozymes were modified to

resemble the *Pneumocystis* ribozyme. These modifications were made to allow direct comparisons between the individual ribozyme constructs and because P9.0 was shown not to be required (19). As a result, the sequences preceding the terminal guanosine are not predicted to form the P9.0 helix. Therefore, to enhance the ωG binding affinity of the *Tetrahymena* and *Candida* ribozymes, one strategy to try might be to modify the 3' sequence back to the native intron sequence. Nevertheless, our results demonstrate that all three constructs, in their current incarnations, effectively catalyze the TES reaction.

Comparison with Previous Results

A previously characterized *Tetrahymena* ribozyme that lacks both the 5' and 3' exon sequences and the RE1 and RE3 sequences is capable of catalyzing a TES-like reaction (143). This highly truncated *Tetrahymena* ribozyme was shown to be capable of binding pseudoknot structures *in trans* resulting in the formation of P1 and P10 helices. For TES ribozymes, the individual ribozyme constructs contain the nucleotides required for the P1 and P10 helices. Unlike the TES reaction, this reaction requires high magnesium concentrations (10-100 mM), high temperatures (55-65 °C), and an exogenous guanosine cofactor. For our studies, TES reactions conducted with the individual ribozymes were optimized at lower magnesium concentrations (10-25 mM) at lower temperatures (44 °C), and do not require an exogenous guanosine cofactor.

Most recently, an engineered twin ribozyme has been shown to mediate the excision of an internal segment from within an RNA substrate (144). In this reaction, excision is mediated by an exchange reaction, consisting of two cleavage and two ligation reactions. In addition, two RNA substrates are required for the twin ribozyme approach, whereas for TES only one substrate is required. The reaction conditions for the twin ribozyme reaction are similar to TES conditions, however, prohibitively long incubation times (8 h) are required as compared to 30-45 min for TES. Furthermore, the product yield for the twin ribozyme-mediated excision reaction is considerably lower (17%) than we get with TES reactions conducted with *Pneumocystis* ribozyme (68%).

Table 5.1. Optimized TES reaction conditions for the *Pneumocystis*, *Tetrahymena*, and *Candida* ribozyme constructs

Optimized TES reactions conditions[a]				TES product yield and observed rate constants for substrate cleavage and exon ligation reactions[b]		
	$MgCl_2$ (mM)	Time (min)	Ribozyme (nM)	TES product (%)	Substrate Cleavage (k_{obs} min^{-1})	Exon Ligation (k_{obs} min^{-1})
Pneumocystis	15	30	100	68.5 ± 4.4	5.0 ± 1.1	3.2 ± 0.1
Tetrahymena	10	45	166	47.5 ± 0.18	0.17 ± 0.05	0.14 ± 0.02
Candida	25	30	75	50.0 ± 1.5	0.31 ± 0.01	0.34 ± 0.03

[a]The individual ribozyme constructs were optimized for magnesium concentration, time, and ribozyme concentration for maximum TES product yields. [b]The observed rate constants, k_{obs}, were obtained for the substrate cleavage and exon ligation reactions under single turnover conditions.

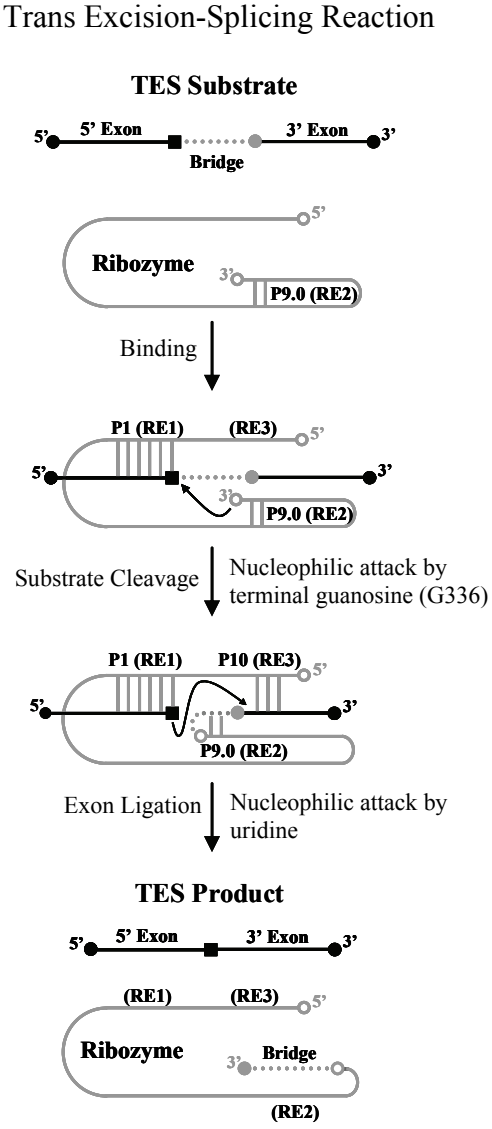


Figure 5.1. Diagram of the ribozyme-mediated trans excision-splicing reaction. The *Pneumocystis carinii* ribozyme is denoted in gray lines and the 5' and 3' exon sequences are denoted by black lines. The black square within the 5' exon region denotes a uridine and the gray circle adjacent to the 3' exon region denotes a guanosine (ωG). The bridge region, to be excised, is denoted as a dotted line. The P1 helix is formed through base-pairing of recognition element 1 (RE1) with the 5' exon. The P10 helix is formed through

base-pairing of recognition element 3 (RE3) with the 3' exon. The first step (substrate cleavage) is catalyzed by G336 (open circle at 3'-end of ribozyme) resulting in the covalent attachment of the 3'-end of the substrate to the 3'-end of the ribozyme. The second step (exon ligation) proceeds through attack of the uridine upon the ωG of the substrate, resulting in ligation of the flanking sequences. Note that the product of the TES reaction is that ωG (of substrate) remains attached to the 3'-end of the ribozyme.

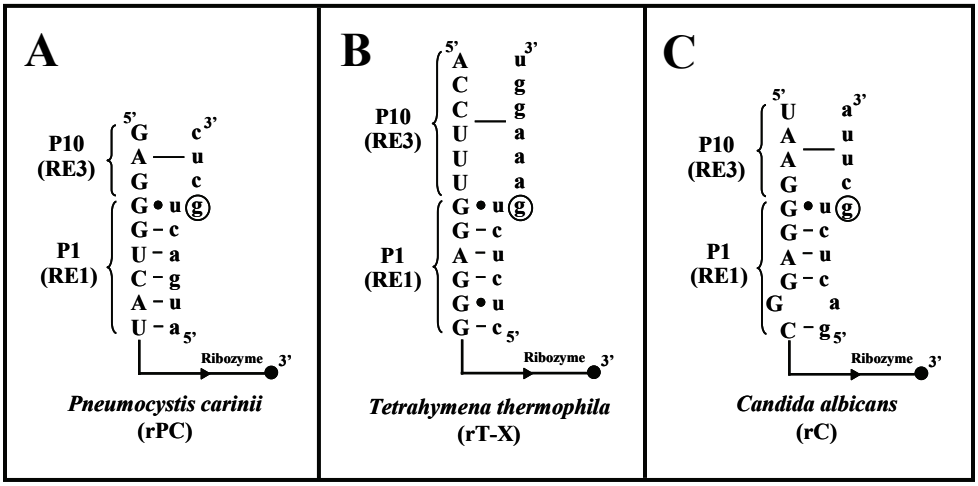


Figure 5.2. Diagram of *Pneumocystis carinii* (rPC), *Tetrahymena thermophila* (rT-X), and *C. albicans* (rC) ribozyme constructs. The sequence *P. carinii*, *T. thermophila*, and *C. albicans* ribozyme constructs are shown in uppercase letters with their respective substrate shown in lowercase letters. Note, only the sequences that are involved with binding the individual substrates are shown. The remaining ribozyme sequences, represented with a line, are different for each of the constructs. The nucleotide to be excised (ωG) is circled for all three ribozyme constructs.

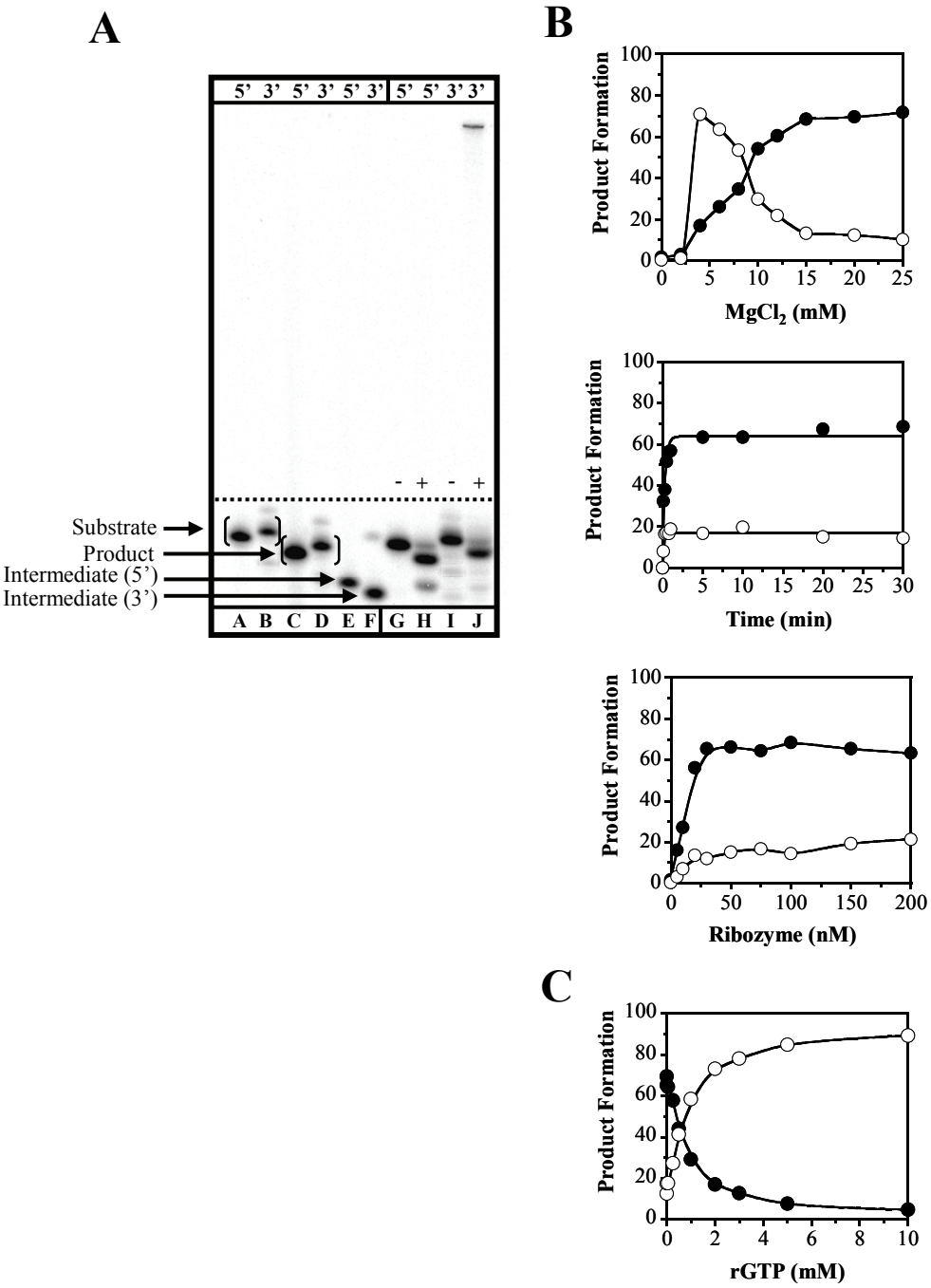


Figure 5.3. TES reactions conducted with the *Pneumocystis carinii* (rPC) ribozyme. A. Polyacrylamide gel representing substrates and products from TES reactions using both 5' and 3'-end radiolabeled substrates using 100 nM rPC, 15 mM $MgCl_2$, at 30 min at 44 °C. Lanes A, C, and E contain 5'-end radiolabeled 10-mer, 9-mer, and 6-mer size

controls. Lanes B, D, and F contain 3'-end radiolabeled 10-mer, 9-mer, and 4-mer size controls. Note that the 3'-end radiolabeled size controls are one nucleotide larger than 5'-end radiolabeled size controls. Gel lanes containing a plus (+) sign show reactions conducted in H15Mg and lanes containing a minus (-) sign show reactions conducted in H0Mg. Lanes G-J contain the normal rPC ribozyme with 5'-end radiolabeled 13-mer (lanes G and H) or 3'-end radiolabeled 13-mer (lanes I and J). B. Graphs of optimization of TES reactions using 5'-end radiolabeled 13-mer substrate for magnesium concentration, time, and ribozyme concentration. All reactions were conducted as above except for the changing variable. C. Graph of rGTP concentration dependence for TES reactions conducted with the rPC ribozyme. All data points represent the average of at least two independent reactions, with standard deviations typically below 10%. Note for all graphs, the 9-mer TES product is represented by filled circles and the substrate cleavage product, 6-mer, is represented as open circles.

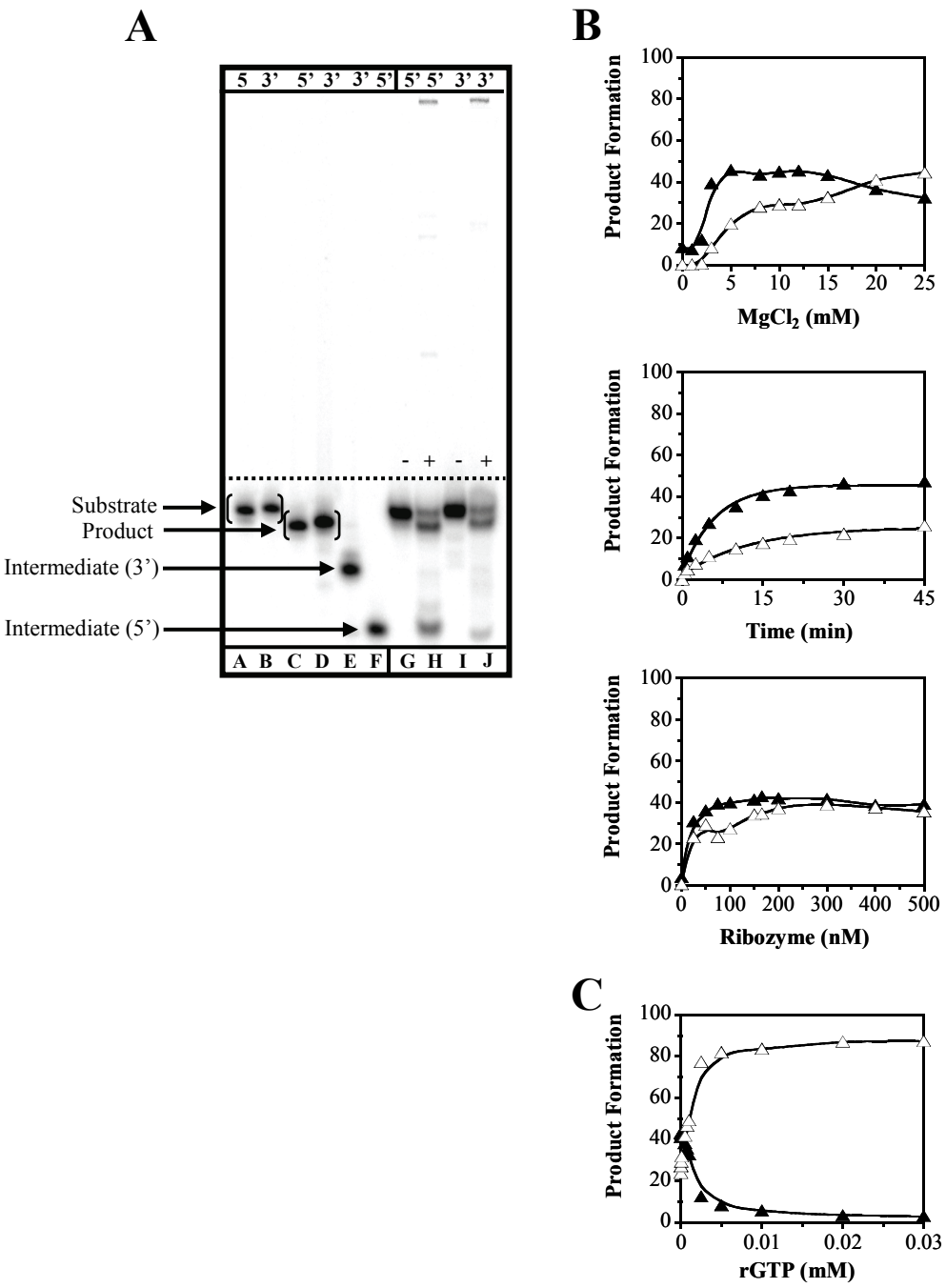


Figure 5.4. TES reactions conducted with the *Tetrahymena thermophila* (rT-X) ribozyme. A. Polyacrylamide gel representing substrates and products from TES reactions using both 5' and 3'-end radiolabeled substrates using 166 nM rT-X, 10 mM $MgCl_2$, at 45 min at 44 °C. Lanes A, C, and F contain 5'-end radiolabeled 13-mer, 12-mer, and 6-mer size

controls. Lanes B, D, and E contain 3'-end radiolabeled 13-mer, 12-mer, and 7-mer size controls. Note that the 3'-end radiolabeled size controls are one nucleotide larger than 5'-end radiolabeled size controls. Gel lanes containing a plus (+) sign show reactions conducted in H10Mg and lanes containing a minus (-) sign show reactions conducted in H0Mg. Lanes G-J contain the normal rT-X ribozyme with 5'-end radiolabeled 10-mer (lanes G and H) or 3'-end radiolabeled 10-mer (lanes I and J). B. Graphs of optimization of TES reactions using 5'-end radiolabeled 13-mer substrate for magnesium concentration, time, and ribozyme concentration. All reactions were conducted as above except for the changing variable. C. Graph of rGTP concentration dependence for TES reactions conducted with the rT-X ribozyme. The data represents the average of at least two independent reactions, with standard deviations typically below 10%. All data points represent the average of at least two independent reactions with standard deviations typically below 10%. Note for all graphs, the 12-mer TES product is represented by filled triangles and the substrate cleavage product, 6-mer, is represented as open triangles.

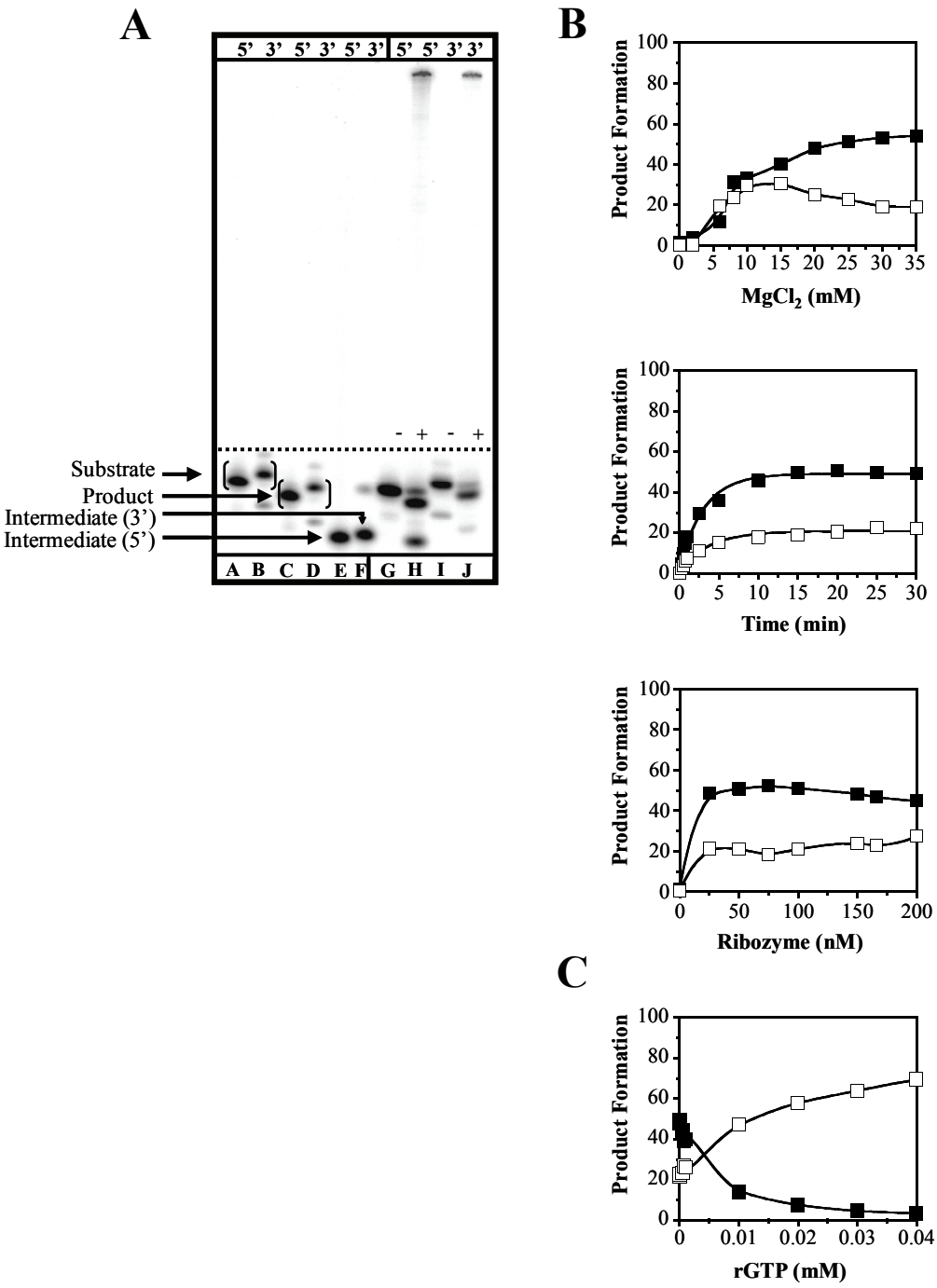


Figure 5.5. TES reactions conducted with the *Candida albicans* (rC) ribozyme. A. Polyacrylamide gel representing substrates and products from TES reactions using both 5' and 3'-end radiolabeled substrates using 75 nM rC, 25 mM $MgCl_2$, at 30 min at 44 °C. Lanes A, C, and E contain 5'-end radiolabeled 11-mer, 10-mer, and 6-mer size controls.

Lanes B, D, and F contain 3'-end radiolabeled 11-mer, 10-mer, and 5-mer size controls. Note that the 3'-end radiolabeled size controls are one nucleotide larger than 5'-end radiolabeled size controls. Gel lanes containing a plus (+) sign show reactions conducted in H25Mg and lanes containing a minus (-) sign show reactions conducted in H0Mg. Lanes G-J contain the normal rC ribozyme with 5'-end radiolabeled 10-mer (lanes G and H) or 3'-end radiolabeled 10-mer (lanes I and J). B. Graphs of optimization of TES reactions using 5'-end radiolabeled 11-mer substrate for magnesium concentration, time, and ribozyme concentration. All reactions were conducted as above except for the changing variable. C. Graph of rGTP concentration dependence for TES reactions conducted with the rC ribozyme. All data points represent the average of at least two independent reactions, with standard deviations typically below 10%. Note for all graphs, the 11-mer TES product is represented by filled triangles and the substrate cleavage product, 6-mer, is represented as open triangles.

CHAPTER SIX-INSERTION OF MODIFIED OLIGORIBONUCLEOTIDES UITLIZING THE TRANS INSERTION-SPLICING REACTION

INTRODUCTION

The site-specific modification of RNA, specifically the incorporation of modified nucleotides, has greatly increased our fundamental understanding of RNA structure and function (145-147). To date, several strategies have been developed for the site-specific alteration of RNA, including the addition of a modified nucleotide(s) (148-157). These include the chemical synthesis of RNA using existing phosphoramidite strategies. In addition, chemical and enzymatic ligation strategies have been developed, which take advantage of the unique properties of the termini of RNA (148-150). Strategies substituting modified nucleosides or using unnatural base-pair combinations during *in vitro* transcription reactions have also been developed (151-155). Furthermore, a twin ribozyme was engineered to replace an internal segment of RNA with a modified nucleotide (156). Most recently, a deoxyribozyme was engineered for the insertion of modified nucleotides (157). In some cases, however, existing strategies suffer from a lack of product specificity, low product yields, long reaction times, and are limited by the types of modified nucleotides that can be incorporated.

The trans insertion-splicing (TIS) reaction, catalyzed by a group I intron-derived ribozyme from *Pneumocystis carinii* (*P. carinii*), was recently developed to sequence-specifically insert a predefined RNA sequence into a target RNA substrate (10). The TIS reaction, which utilizes two separate RNA substrates, has been proposed to proceed through three separate chemical steps. These include two consecutive substrate cleavage steps and a single exon ligation step (Figure 6.1).

To increase the general applicability of the TIS reaction, TIS was investigated as a new strategy for the insertion of modified oligonucleotides. Four modified oligonucleotides were designed to test modifications at either the 2' position of the sugar ring (deoxy), the phosphodiester backbone (phosphothioate), or the base (2-aminopurine or 4-thiouridine) position. All four modified oligonucleotides were shown to be inserted sequence-specifically within their intended target. In addition, a doubly modified

oligonucleotide, including a methoxy and deoxy substitution, was shown to be an effective substrate for the insertion of multiple modified oligonucleotides. The results demonstrate TIS as an alternative strategy for the insertion of modified nucleotides into RNA, including multiple modified nucleotides. The TIS strategy could potentially lead to a more effective method for the creation of full-length RNA transcripts containing these site-specific modifications.

MATERIALS AND METHODS

Oligonucleotide Synthesis and Preparation

DNA oligonucleotides were obtained from Integrated DNA Technologies (Coralville, IA) and were used without further purification. RNA oligonucleotides were obtained from Dharmacon Research, Inc. (Lafayette, CO) and were deprotected using the manufacturer's standard protocol. Concentrations of individual oligonucleotides were calculated from UV absorption measurements using a Beckman DU 650 spectrophotometer (Beckman Coulter, Inc.; Fullerton, CA). RNA oligonucleotides were 5'-end radiolabeled by phosphorylation of the 5' terminal hydroxyl group with [γ-^{32}P] ATP (Amersham Pharmacia Biotech; Piscataway, NJ) using T4 polynucleotide kinase (New England Biolabs; Beverly, MA) as previously described (75). The sequences of the substrates used are shown in Table 6.1.

Ribozyme Synthesis

The previously constructed PC plasmid (8) was linearized in a 50 µL reaction consisting of 16 µg of plasmid, 1 X REACT 2 buffer, and 50 units *Xba*I (Invitrogen) at 37 °C for 2 h. The resulting linearized plasmid was isolated using the QIAquick PCR purification kit (Qiagen) using the manufacturers protocol. Run-off transcription was performed for 2 h in a 50 µL reaction consisting of 1 µg linearized DNA template, 50 units of T7 RNA polymerase (New England Biolabs; Beverly, MA), 40 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 5 mM DTT, 5 mM spermidine, and 1 mM rNTPs. The rPC ribozyme was purified using a Qiagen Plasmid Midi Kit (Qiagen), as previously described (8). Ribozyme concentrations were calculated from UV absorption

measurements using a Beckman DU 650 spectrophotometer (Beckman Coulter, Inc).

TIS Reactions

TIS reactions were conducted under previously optimized reaction conditions using the rPC ribozyme (9). Briefly, a 5.0 μL ribozyme solution consisting of 240 nM rPC in H10Mg (50 mM HEPES (25 mM Na^+). 135 mM KCl, and 10 mM $MgCl_2$ at pH 7.5) was pre-incubated at 60 °C for 5 min followed by slow cooling to 44 °C. A separate solution consisting of 6 nM 5'-end radiolabeled TIS substrate ($^{5'}$AUGACUAAACAU$^{3'}$) and 30 μM of the individual TIS inserts ($^{5'}$GCUCU<u>CGU</u>G$^{3'}$, underlined nucleotides denote the sites of modification) in H10Mg was pre-incubated at 44 °C. TIS reactions were initiated by the addition of 1.0 μL of the precombined TIS donor and acceptor substrates to the ribozyme solution. This results in a final concentration of 200 nM rPC ribozyme, 1.0 nM TIS acceptor substrate, and 1.0 μM TIS donor substrate. TIS reactions were conducted for 2 h at 44 °C, at which time the reaction was terminated by addition of an equal volume (6 μL) of a 2X stop buffer (10 M urea, 3 mM EDTA, and 0.1X TBE). The reaction mixture was denatured for 1 min at 90 °C and separated on a 12% polyacrylamide/8 M urea gel. The gel was transferred to chromatography paper and dried under vacuum. The bands were visualized and quantified on a Molecular Dynamics Storm 860 Phosphorimager.

RESULTS

Design of the TIS Test System

It has previously been shown that a group I intron-derived ribozyme from *P. carinii* (rPC) can catalyze the TIS reaction, as shown in Figure 6.1 (9). For these previous studies, the sequence of the ribozyme was kept intact, albeit truncated, and the donor and acceptor substrates closely mimicked the ribozymes native 5' and 3' exon sequences. Maintaining these sequences was important in order to maximize the binding and subsequent reactivity of the tri-component system during the original proof-of-concept study (10). Therefore, the same ribozyme and substrate sequences were used in this study, except for the nucleotide analog substitutions at positions 6-8 (Figure 6.1) within

the TIS donor substrate (see Table 6.1). Though positions 6-8 are involved in binding and subsequent alignment of TIS substrates, through molecular interactions with the ribozyme, these sites were chosen as they would be expected to have the least deleterious effect upon the TIS reaction (Figure 6.2). To determine the generality of the method, nucleotides that contain sugar, phosphate, and nucleobase modifications were analyzed. These include deoxyribose (dG), methoxyribose (mU), phosphorothioate (SH), 2-aminopurine (2AP), and 4-thiouridine (4SU) substitutions (Table 6.1). Note that for each of the donor substrates, the same sized TIS product is expected to form, which is the same size as that with the previously sequenced unmodified substrate (10).

TIS Insertion of Modified Oligoribonucleotides

TIS reactions were conducted as previously described (10), under previously optimized reaction conditions (see Figure 6.3 legend). Substrates, intermediates, and products of each of the individual TIS reactions were separated and visualized via polyacrylamide gel electrophoresis. The initial reactions focused on modifications only at position 7, as this was expected to be least disruptive to the reaction. As shown in Figure 6.3, each of the four substrates modified at position 7 resulted in the expected 18-mer product band (Lanes F-J). Note that the relatively small size differences between the TIS products are due to different modifications within the substrates. Apparently, position 7 of the donor substrate is readily modifiable; resulting in the effective insertion of modified RNAs into other RNAs (see reaction yields in Table 6.2). At position 7, the deoxy (Figure 6.3, lane G) and phosphothioate (Figure 6.3, lane H) substitutions were essentially as effective as the non-modified substrate (Figure 6.3, lane F). TIS product yields for the deoxy was 72 ± 0.8% and 65 ± 0.2% for the phosphothioate, compared to 68 ± 1.8% for the non-modified substrate (Table 6.2). Lower, although comparable, yields (59 ± 0.9% versus 63 ± 0.3%) were obtained for the two nucleobase modifications (Figure 6.3, lanes I and J). This result suggests that base modifications do hinder TIS product formation, most probably through interrupting structural elements within the rPC ribozyme utilized during the course of the reaction. Note that the faint second product band for TIS reactions conducted with the 4-thiouridine TIS donor substrate (Figure 6.3, lane J) is potentially due to degradation of the starting material.

It was then tested whether more than one modification could be added to the donor substrate. Positions 6 and 8 were analyzed to see if they too could be modified (in addition to position 7), and different modifications, including a methoxy substitution and a deoxy substitution, were chosen to broaden the potential applicability of the method. As shown in Figure 6.3 (Lane K), the doubly modified donor substrate is as effective a TIS substrate (75 ± 0.9% product) as its non-modified counterpart. Apparently, the 2' position of the ribose at positions 6 and 8 (in addition to position 7) of the donor substrate are sites that can be readily modified. This result shows that the TIS reaction is an effective strategy for the insertion of multiple modifications within an RNA transcript. Taken together, these results demonstrate that the TIS reaction can be exploited for inserting non-native, chemically modified oligonucleotides into RNAs *in trans*.

DISCUSSION

In this study, it has been demonstrated that the trans insertion-splicing ribozyme from *P. carinii* can utilize substrates that contain modifications, and that this activity can be exploited to generate RNAs with one or more internal modifications. In this study, the functional group modifications within the sugar phosphate backbone at position 7 of the donor substrate (see Figure 6.2) have been shown to not hinder the TIS reaction. In contrast, nucleobase modifications at position 7 do reduce TIS yields, although not by prohibitive amounts (approximately 15% reduction). These results suggest that disruption of the P9.0 helix interaction between this nucleobase and the ribozyme during the second TIS reaction step is not prohibitive (see Figure 6.1). Therefore, the TIS reaction is a relatively effective method for synthesizing RNAs that contain sugar-phosphate backbone and nucleobase modifications.

Like that at position 7, ribose modifications at donor substrate positions 6 and 8 also do not reduce TIS yields. Apparently, the 2' OH groups at positions 6, 7, and 8 are expendable for the TIS reaction, which enhances the flexibility of using TIS as a synthetic tool. In addition, that the TIS reaction conducted with the doubly-modified donor substrate was successful demonstrates that TIS is an effective strategy for the insertion of multiple modified nucleotides within a given RNA sequence.

While the products of the TIS reactions in this proof-of-concept study are short enough to have been synthesized chemically in high yields, chemically synthesizing large RNAs, with or without modifications, is prohibitive due to insufficient expected yields. It is envisioned that the TIS reaction could be exploited for the synthesis of large, site-specifically modified RNAs. This could be accomplished using TIS to insert a small, chemically synthesized RNA (acting as the donor molecule, modified at a position corresponding to positions 6, 7, or 8 in Figure 6.1) into a full-length RNA transcript (acting as the acceptor substrate). Note that in terms of ribozyme design, all of the ribozyme and substrate sequences that make up the molecular recognition elements can be altered as desired to create appropriate target-substrate combinations, as long as the substrate-ribozyme base pairs shown in Figure 6.1 are maintained. In addition, it is expected that full-length transcripts could be effective TIS acceptor substrates because a similar ribozyme, acting as a trans excision-splicing (TES) catalyst, has already been shown to be active on a full-length GFP transcript *in vivo* (10).

Comparison with Previously Developed Methods

Enzymatic methods have previously been developed for the incorporation of modified nucleotides into large RNAs. One strategy is to add the desired modified nucleotide precursor directly into transcription reactions conducted *in vitro* (151, 152). Although this strategy can be quite effective and useful, it is not appropriate for synthesizing single site-specific RNAs, as the modified nucleotides are incorporated throughout the entire transcript. Another strategy solves this problem by utilizing unique non-native nucleotides in the transcription template that only pair with other unique RNA nucleotides, thus forming RNAs during transcription that have non-native nucleotides incorporated (153-155). Unfortunately, the numbers of non-native nucleotides that can be used in this strategy is extremely limiting. In comparison, the TIS strategy could potentially be simpler to design and enact, and is not entirely dependent on the secondary structure formation properties of the modified nucleotide(s) for the success of the method.

The enzymatic ligation of separately synthesized RNAs, including RNAs that contain modifications, using either RNA or DNA ligases (149, 150, 158-161) have been

the most widely used methods for constructing large modified RNAs. Though these methods have been quite successful, in some instances, they still suffer from low product yields and long reaction times. In addition, the substrates that are used in enzymatic ligation methods are typically transcript length, which requires highly precise synthetic, purification, and ligation methods. In contrast, the TIS method requires the ribozyme (whose sequence has been designed for the particular application), one transcript, and a synthetic modified RNA.

Recently, the site-specific modification of RNA has been accomplished using an engineered dual hairpin ribozyme, termed a twin ribozyme (156). In this strategy, the ribozyme can replace an internal segment of an RNA transcript with a separate RNA which may or may not contain a modification. This method shares many similarities to the TIS reaction, including using a catalytic RNA and two exogenous RNA substrates. Nevertheless, the two methods differ by the type of the ribozyme (hairpin versus group I introns), the required sequence elements in the substrates, the catalytic mechanisms, and the optimal reaction conditions. The TIS strategy, however, results in higher product yields than the twin ribozyme approach (70% versus 20% product yield), which suggests that the TIS strategy could be more efficient at inserting modified oligonucleotides into RNA transcripts (156).

Table 6.1. TIS starting material and insert substrate sequences.

TIS Acceptor Substrate	Sequence
12-mer-starting material	$^{5'}$AUGACUAAAGAU$^{3'}$
Unmodified TIS Donor Substrate	**Sequence**
9-mer-unmodified insert	$^{5'}$GCUCUCGUG$^{3'}$
Modified TIS Donor Substrate	**Sequence**
9-mer-deoxyguanosine (dG)	$^{5'}$GGUCUC(dG)UG$^{3'}$
9-mer-phosphothioate (SH)	$^{5'}$GGUCUC(SH)GUG$^{3'}$
9-mer-2-aminopurine (2AP)	$^{5'}$GGUCUC(2AP)UG$^{3'}$
9-mer-4-thiouridine (4SU)	$^{5'}$GGUCUC(4SU)UG$^{3'}$
9-mer-doubly modified (DM)	$^{5'}$GGUCU(dC)G(mU)G$^{3'}$

The sequences for the TIS acceptor substrate (12-mer), normal (unmodified) TIS donor substrate (9-mer), and modified TIS donor substrates (9-mer) are shown. For both the unmodified and modified TIS donor substrates the sequences to be inserted are highlighted in gray. For each modified TIS donor substrate the modified nucleotide (located at either position C6, G7, or U8) is shown in parentheses. The modified TIS donor substrates include a deoxyguanosine or deoxycytosine (dG or dC), phosphothioate (SH), 2-aminopurine (2AP), 4-thiouridine (4SU), or methoxyuridine (mU) substitution.

Table 6.2. Percent TIS product formation for individual modified oligonucleotides.

TIS Donor Substrate	Predicted TIS Product Sequence[a] (18-mer)	TIS Product[b]
9-mer-normal insert	$^{5'}$AUGACUCUCGUGAAACAU$^{3'}$	68±1.8%
9-mer-deoxyguanosine (dG)	$^{5'}$AUGACUCUC(dG)UGAAACAU$^{3'}$	72±0.8%
9-mer-phosphothioate (SH)	$^{5'}$AUGACUCUC(SH)GUGAAACAU$^{3'}$	65±0.2%
9-mer-2-aminopurine (2AP)	$^{5'}$AUGACUCUC(2AP)UGAAACAU$^{3'}$	59±0.9%
9-mer-4-thiouridine (4SU)	$^{5'}$AUGACUCUC(4SU)UGAAACAU$^{3'}$	63±0.3%
9-mer-doubly modified (DM)	$^{5'}$AUGACUCU(dC)G(mU)GAAACAU$^{3'}$	75±0.9%

[a]Predicted TIS product (18-mer) for TIS reactions conducted with the TIS acceptor substrate (12-mer) and either the unmodified or modified TIS donor substrates. The sequences highlighted in gray represent the insert region including the modified nucleotide(s). TIS reactions were conducted under previously optimized reaction conditions (200 nM ribozyme, 10 mM $MgCl_2$, 1 μM insert at 44 °C for 2 hr). [b]Percent TIS product (18-mer) formed using each of the modified oligonucleotide substrates. The TIS product, $^{5'}$AUGACUCUCGUGAAACAU$^{3'}$, includes modification at either the C6, G7, or U8 positions (shown in parentheses) from the TIS donor substrate.

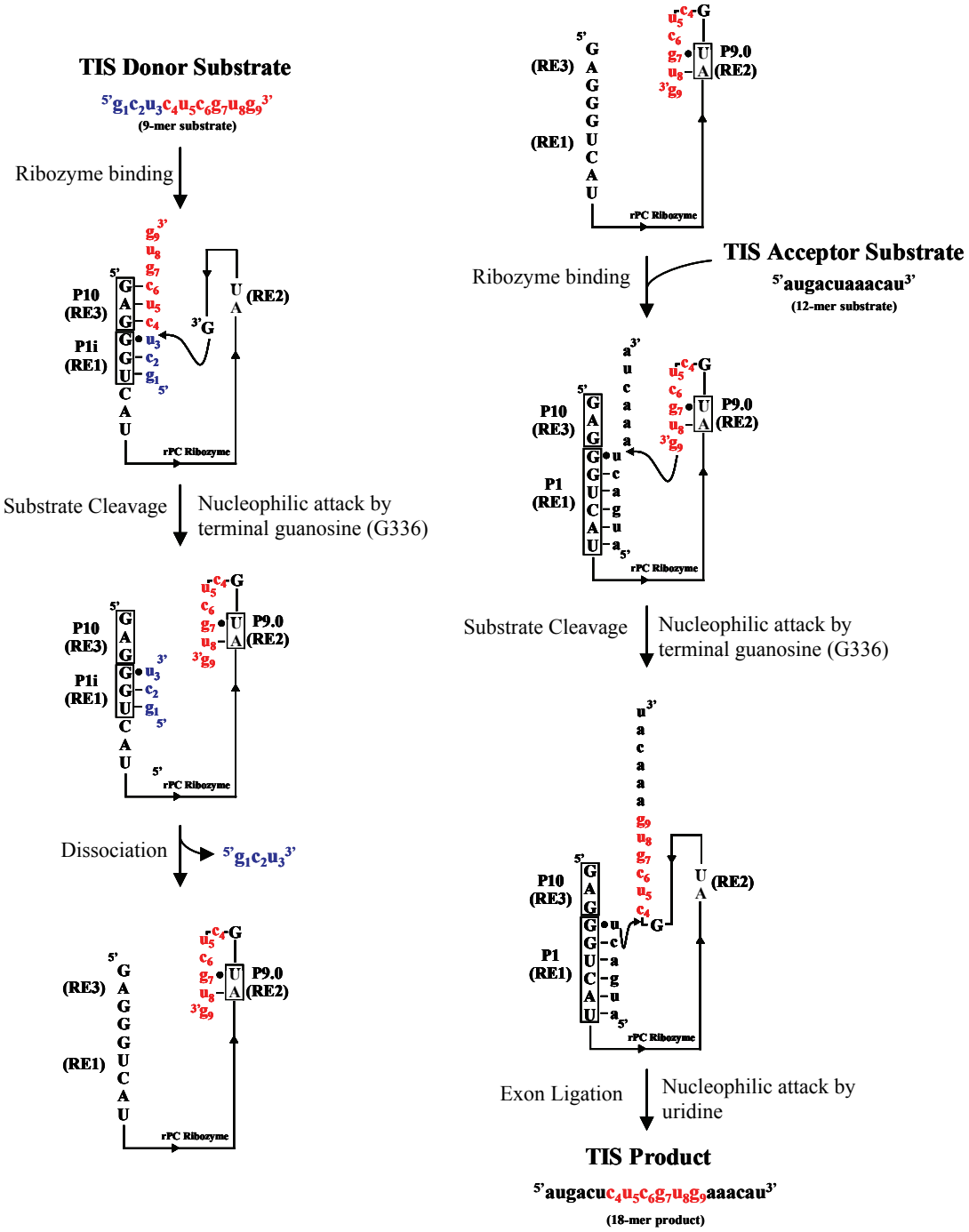


Figure 6.1. Schematic representation of the TIS reaction model system. The *Pneumocystis carinii* (rPC) ribozyme is shown in black uppercase lettering including the ribozyme recognition elements; RE1, RE2, and RE3 that form the P1, P9.0, and P10 helices respectively. The 9-mer TIS donor substrate ($^{5'}G_1C_2U_3C_4U_5C_6G_7U_8G_9{}^{3'}$) is shown

in both blue lettering (5’ exon region) and red lettering (insert region). The guanosine at the 3’ terminal end of the insert region represents ωGi. The 12-mer TIS acceptor substrate ($^{5'}$AUGACUAAACAU$^{3'}$) is shown in black lowercase lettering. The TIS donor substrate forms the P1i and P10 helices through RE1 and RE3 of the ribozyme. The 3’ terminal guanosine (G336) of the ribozyme performs a nucleophilic attack upon the 5’ splice site, resulting in the covalent attachment of the insert region to the 3’-end of the ribozyme. The 5’ exon region of the TIS donor substrate dissociates from the ribozyme, and the TIS acceptor substrate is bound through a second P1 helix. The ωGi (aligned by P9.0 helix formation) performs a nucleophilic attack upon the 5’ splice site in the TIS acceptor substrate, resulting in the covalent attachment of the 3’ exon region of the TIS acceptor substrate to the 3’-end of the ribozyme. Exon ligation proceeds through the nucleophilic attack of the uridine at the 3’-end of 5’ exon region of the TIS substrate upon G336 of the ribozyme, resulting in the final TIS product (18-mer, $^{5'}$AUGACUC$_4$U$_5$C$_6$G$_7$U$_8$G$_9$AAACAU$^{3'}$).

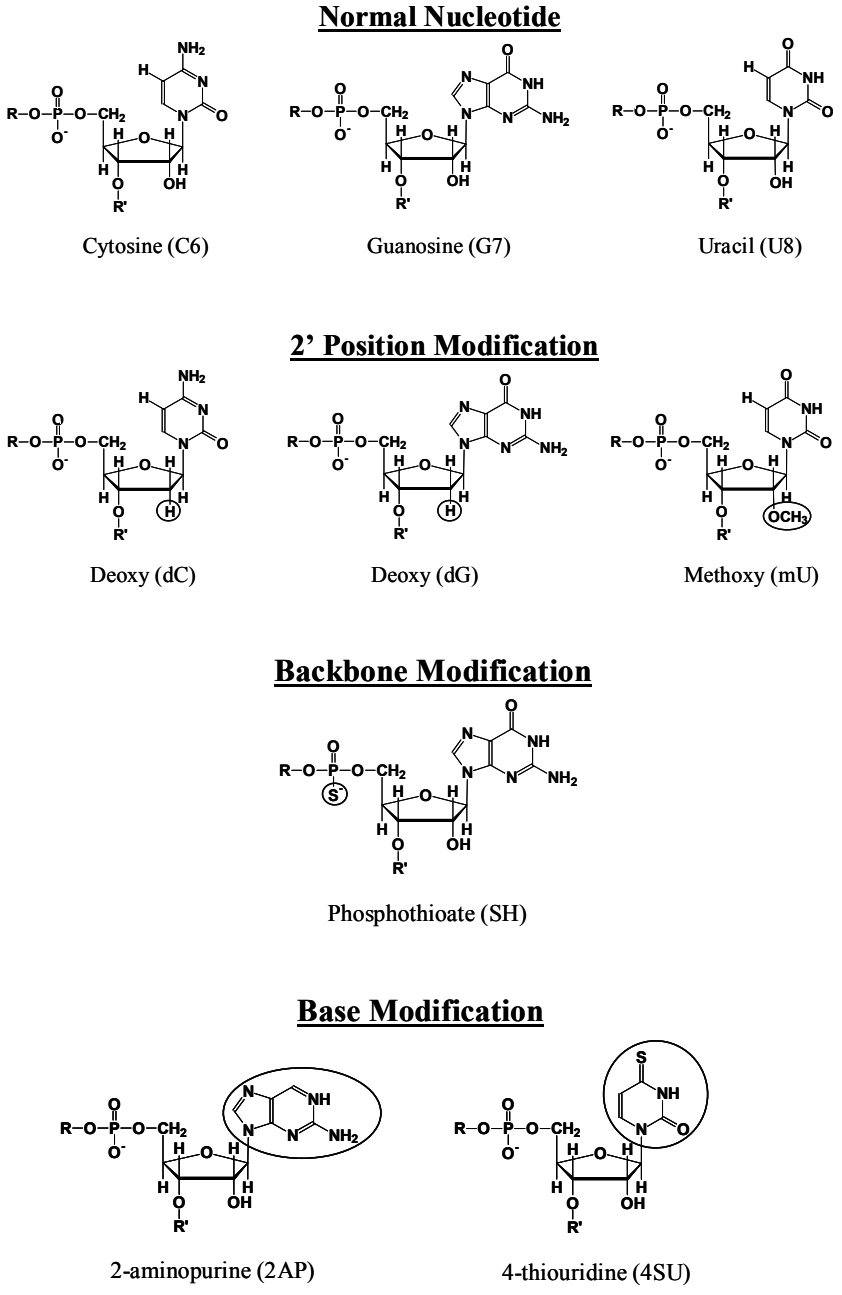


Figure 6.2. Diagram of the modified nucleotides and naming conventions used in this study. The normal nucleotide at position 6, 7, and 8 within the TIS donor substrate (9-mer, $^{5'}G_1C_2U_3C_4U_5C_6G_7U_8G_9{}^{3'}$) are shown in addition to the individual modifications to the 2' position of the sugar ring [deoxycytosine (dC), deoxyguanosine (dG), and methoxyuridine (mU)], the phosphodiester backbone [phosphothioate (SH)], and base

position [2-aminopurine (2AP) and 4-thiouridine (4SU)]. Note each individual modification is depicted with a circle.

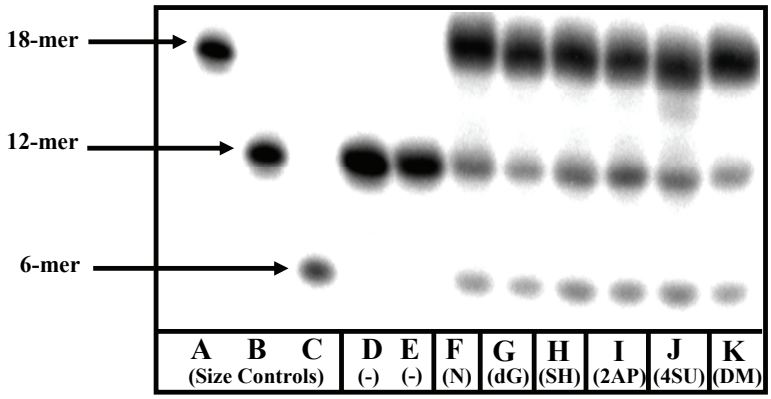


Figure 6.3. Polyacrylamide gel showing substrates, intermediates, and products of the trans insertion-splicing reaction. Reactions were conducted with 200 nM ribozyme, 1 nM acceptor substrate, and 1 μM donor substrate in H10Mg at 44 °C for 2 h. Lanes A, B, and C contain 5'-end radiolabeled 18-mer, 12-mer, and 6-mer size controls respectively. Negative control lanes (-) consist of reactions run in H0Mg buffer (Lane D) or without rPC ribozyme (Lane E). Lane F contains the TIS reaction with the unmodified donor substrate (N), and serves as a positive control. Lanes G, H, I, J, and K contain the TIS reaction with deoxy (dG), phosphothioate (SH), 2-aminopurine (2AP), 4-thiouridine (4SU), and doubly-modified (DM) substituted donor substrates, respectively.

CHAPTER SEVEN-FUTURE DIRECTIONS

The Trans Excision-Splicing Reaction

The trans excision-splicing (TES) reaction has been envisioned as a potential therapeutic strategy for the correction of defective RNA transcripts within a cell. In this strategy, the *Pneumocystis carinii* ribozyme would be engineered to excise specific regions from within defective RNA transcripts. The resulting corrected version of the transcript will lead to the production of a functional protein. As seen for other ribozyme-mediated reactions, similar to the TES reaction, repairing as little as 10% of these transcripts has beneficial effects for an underlying genetic mutation (85, 88). The work presented herein has answered some of the questions relating to the TES reaction with respect to the overall reaction mechanism.

Intramolecular Transesterification Reaction for Cleavage of TES Substrates

The work presented herein has established a mechanistic framework for the substrate cleavage step of the TES reaction. Substrate cleavage was shown to be a ribozyme-mediated intramolecular transesterification reaction, catalyzed by the 3' terminal guanosine (G336) of the *Pneumocystis carinii* ribozyme. The direct consequence of the intramolecular transesterification reaction is that the bridge-3' exon region of TES substrates will be covalently attached to the 3'-end of the *Pneumocystis* ribozyme. Proof of the mechanism was obtained through sequencing of ribozyme reaction intermediates isolated from *in vitro* and *in vivo* TES reactions. The results obtained confirmed the presence of the expected ribozyme intermediate.

Though the mechanism for the substrate cleavage step of TES was established, several questions arise with respect to enhancing the reactions specificity and efficiency *in vitro* and *in vivo*. For instance, it is not known how effectively the *Pneumocystis* ribozyme terminates in the required 3' terminal guanosine. As future TES reactions will undoubtedly be conducted in eukaryotic cell lines, enhancing the ribozymes ability to terminate in a guanosine is critical. For instance, a *Pneumocystis* ribozyme, expressed within a eukaryotic cell line (HeLa), lacking a 3' terminal guanosine resulted in a ribozyme incapable of catalyzing the TES reaction. To address this concern, the

percentage of *Pneumocystis* ribozymes that terminate in guanosine could be obtained through 3' rapid amplification of cDNA ends (RACE). This strategy could be used both *in vitro* and *in vivo* to distinguish between active and inactive forms of the *Pneumocystis* ribozyme.

To date, it is unknown how effectively the 3' terminal guanosine (G336) of the *Pneumocystis* ribozyme is aligned into the guanosine binding site (GBS). If the terminal guanosine is found to be poorly aligned into the GBS, the molecular recognition elements of the ribozyme could be modified to enhance this alignment. Specifically, the two nucleotides, which constitute recognition element 2 (RE2), could be altered to base-pair with the two nucleotides preceding G336 of the ribozyme. This would lead to a P9.0 helix interaction, which has been shown to aid in the alignment of the 3' terminal guanosine for other ribozyme-mediated reactions (17).

If intramolecular 3' hydrolysis is shown to be inefficient at creating the active form of the *Pneumocystis* ribozyme *in vivo*, the ribozyme could potentially be reengineered. This would involve including both a P9.0 and P10 helix interaction, which could potentially increase the efficiency of the hydrolysis reaction. As *Pneumocystis* ribozymes expressed *in vivo* will likely have transcription termination sequences defining their 3'-ends, enhancing the intramolecular 3' hydrolysis reaction will increase the likelihood of successful TES reactions.

Lastly, it has been shown that the substrate cleavage reaction *in vivo* is not highly sequence-specific. For example, the *Pneumocystis* ribozyme was previously shown *in vivo* to be capable of cleaving a second site within its target RNA transcript. To address this concern, the *Pneumocystis* ribozyme could be reengineered to include an extended guide sequence (83, 84, 86-90) to aid in enhancing the binding of the ribozyme to the intended RNA transcript. Though there are many questions still to be answered for the TES reaction, the mechanistic investigation presented in this work has laid a foundation for future studies for improving the overall TES reaction *in vitro* and *in vivo*.

Kinetic Characterization of the Substrate Cleavage Reaction of the TES Reaction

The work presented herein has established a kinetic framework for the substrate cleavage step of the TES reaction. Several conclusions can be drawn from the kinetic

framework, including that the rate constant for the substrate cleavage reaction is approximately 60-fold lower than a previously reported *Tetrahymena thermophila* ribozyme. The rate constant for substrate cleavage is only 4-fold greater than that for substrate dissociation. Furthermore, the rate of product dissociation is slower than the rate of substrate cleavage, which curtails the possibility of multiple-turnover reactions. Lastly, the results indicate that a conformational change exists between the two steps of the TES reaction.

The kinetic study has been informative for increasing our understanding of the fine details by which the substrate cleavage step proceeds. The conclusions that can be drawn from the kinetic data, however, suggest strategies whereby we could enhance the rate and yield of the substrate cleavage reaction. For example, the relatively slow rate of product dissociation is potentially rate-limiting for any potential multiple-turnover reactions. To increase the rate of product dissociation, the ribozyme could be reengineered to weaken the binding of TES products, for example by shortening helix P1. A shortened P1 helix has previously been shown to increase the rate of product dissociation for a group I intron-derived ribozyme from *Anabaena* (72).

Separately, the rate of substrate dissociation was shown to be comparable to the rate of substrate cleavage. This suggests that a substantial amount of substrate will dissociate prior to the substrate cleavage reaction. To slow the rate of substrate dissociation, the ribozyme could be reengineered to strengthen the binding of TES substrates, for example by elongating helix P1. Elongating helix P1, however, could strengthen the binding of TES products, which could decrease the rate of product dissociation even further. Therefore, TES ribozymes would have to be designed to include a P1 helix, of optimal length, to overcome the limitations set by the rates for substrate and product dissociation. Lastly, the length of helix P1 would have to investigated to determine the effects upon the specificity of the TES reaction.

Group I Intron-Derived Ribozyme Catalyze the Trans Excision-Splicing Reaction

The work presented herein has demonstrated that the TES reaction is generally applicable to other group I intron-derived ribozymes. A group I intron-derived ribozyme from *Tetrahymena thermophila* and *Candida albicans* were shown to catalyze the

excision of a single nucleotide from within their respective RNA substrate. Moreover, both ribozyme constructs catalyze the substrate cleavage step through an intramolecular transesterification reaction, similar to the *Pneumocystis carinii* ribozyme. The *Tetrahymena* and *Candida* ribozymes, however, were shown to catalyze the individual steps of the TES reaction less efficiently than the *Pneumocystis* ribozyme. The decrease in efficiency was attributed to the *Tetrahymena* and *Candida* ribozymes inability to effectively bind their terminal guanosine into the GBS. To increase the ribozymes affinity for their terminal guanosine, the ribozymes' RE2 recognition elements, could be modified (to form helix P9.0) to aid in aligning their terminal guanosines.

That additional group I intron-derived ribozymes can catalyze the TES reaction, raises questions with respect to the role of these ribozymes *in vivo*. Group I intron-derived ribozymes are essentially exon-less introns that mimic the final intron (L-IVS) product after self-splicing occurs. It has previously been shown that these spliced-out introns undergo intramolecular cyclization reactions, resulting in a predicted inactive, circular intron product. However, it is not known if these spliced-out introns can catalyze additional intramolecular transesterification reactions. This could include spliced-out introns catalyzing the TES reaction using additional RNA transcripts as substrates. If, however, these spliced-out introns catalyze the cleavage of RNA transcripts that code for essential protein products, this could have deleterious effects on a cell. With this in mind, the intramolecular cyclization reaction could have evolved to deactivate spliced-out introns from cleaving unintended RNA transcripts.

The Trans Insertion-Splicing Reaction

The trans insertion-splicing (TIS) reaction has been envisioned as a new strategy for the construction of RNA transcripts with site-specific modifications. In this strategy, the *Pneumocystis carinii* ribozyme would be engineered to insert specific regions from within full-length RNA transcripts. In turn, these modified RNA transcripts could be useful for functional as well as structural studies of RNA *in vitro* and *in vivo*. The work presented herein has answered some of the questions relating to the feasibility of using the TIS reaction as a strategy for site-specific incorporation of modified oligonucleotides into RNA.

The Trans Insertion-Splicing Reaction can Catalyze the Insertion of Modified Oligonucleotides

The work presented herein has demonstrated that the TIS reaction is a viable strategy for the insertion of modified oligonucleotides into an RNA substrate. In the current study, key positions within a nucleotide were altered, including changes to the 2' position of the sugar ring, the phosphodiester backbone, and the base positions. Modifications at each of these positions can be used as probes for investigating the underlying intricacies of a specific RNAs structure and function. For example, the TIS strategy could be useful for the insertion of spin labels for electron paramagnetic resonance (EPR) or for labeling RNA sequences with relevant isotopes for nuclear magnetic resonance (NMR) studies. In addition, TIS could be used for labeling RNA sequences with fluorescent probes for use in fluorescent resonance energy transfer (FRET) experiments. Finally, the role of post-transcriptionally modified RNA transcripts could be more thoroughly investigated through creation of these RNA transcripts using the TIS strategy. With the broad range of modified nucleotides that are commercially available, future TIS reactions could potentially open new avenues for the study of RNA.

REFERENCE

1. Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., and Cech, T. R. (1982) Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena, *Cell 31*, 147-157.
2. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme, *Cell 35*, 849-857.
3. Cech, T. R. (1990) Self-splicing of group I introns, *Annual Review of Biochemistry 59*, 543-568.
4. Symons, R. H. (1992) Small catalytic RNAs, *Annual Review of Biochemistry 61*, 641-671.
5. Teixeira, A., Tahiri-Alaoui, A., West, S., Thomas, B., Ramadass, A., Martianov, I., Dye, M., James, W., Proudfoot, N. J., and Akoulitchev, A. (2004) Autocatalytic RNA cleavage in the human beta-globin pre-mRNA promotes transcription termination, *Nature 432*, 526-530.
6. Dotson II, P. P., and Testa, S. M. (2006) Group I intron-derived ribozyme recombination reactions, *Recent Developments in Nucleic Acids Research 2*, 307-324.
7. Salehi-Ashtiani, K., Luptak, A., Litovchick, A., and Szostak, J. W. (2006) A genomewide search for ribozymes reveals an HDV-like sequence in the human CPEB3 gene, *Science 313*, 1788-1792.
8. Bell, M. A., Johnson, A. K., and Testa, S. M. (2002) Ribozyme-catalyzed excision of targeted sequences from within RNAs, *Biochemistry 41*, 15327-15333.
9. Baum, D. A., and Testa, S. M. (2005) In vivo excision of a single targeted nucleotide from an mRNA by a trans excision-splicing ribozyme, *RNA 11*, 897-905.
10. Johnson, A. J., Sinha, J., and Testa, S. M. (2005) Trans insertion-splicing: Ribozyme-catalyzed insertion of targeted sequences into RNAs, *Biochemistry 44*, 10702-10710.
11. Cech, T. R., Zaug, A. J., and Grabowski, P. J. (1981) *In vitro* splicing of the ribosomal RNA precursor of Tetrahymena: Involvement of a guanosine nucleotide in the excision of the intervening sequence, *Cell 27*, 487-496.
12. Inoue, T., Sullivan, F. X., and Cech, T. R. (1986) New reactions of the ribosomal RNA precursor of *Tetrahymena* and the mechanism of self-splicing, *Journal of Molecular Biology 189*, 143-165.
13. Zaug, A. J., and Cech, T. R. (1986) The *Tetrahymena* intervening sequence ribonucleic acid enzyme is a phosphotransferase and an acid phosphatase, *Biochemistry 25*, 4478-4482.
14. Zaug, A. J., Been, M. D., and Cech, T. R. (1986) The *Tetrahymena* ribozyme acts like an RNA restriction endonuclease, *Nature 324*, 429-433.
15. Zaug, A. J., and Cech, T. R. (1986) The intervening sequence RNA of *Tetrahymena* is an enzyme, *Science 231*, 470-475.

16. Herschlag, D., and Cech, T. R. (1990) Catalysis of RNA cleavage by the *Tetrahymena thermophila* ribozyme. 1. Kinetic description of the reaction of an RNA substrate complementary to the active site, *Biochemistry 29*, 10159-10171.
17. Mei, R., and Herschlag, D. (1996) Mechanistic investigations of a ribozyme derived from the Tetrahymena group I intron: Insights into catalysis and the second step of self-splicing, *Biochemistry 35*, 5796-5809.
18. Bell, M. A., Sinha, J., Johnson, A. J., and Testa, S. M. (2004) Enhancing the second step of the trans excision-splicing reaction of a group I ribozyme by exploiting P9.0 and P10 for intermolecular recognition, *Biochemistry 43*, 4323-4331.
19. Baum, D. A., Sinha, J., and Testa, S. M. (2005) Molecular recognition in a trans excision-splicing ribozyme: Non-watson-crick base pairs at the 5' splice site and ωG at the 3' splice site can play a role in determining the binding register of reaction substrates, *Biochemistry 44*, 1067-77.
20. Crick, F. H. (1958) On protein synthesis, *Symposia of the Society for Experimental Biology*, *The Biological Replication of Macromolecules*, *XII 12*, 138-163.
21. Crick, F. (1970) Central dogma of molecular biology, *Nature 227*, 561-563.
22. Voet, D., Voet, J.G., Pratt, C.W. (2001) Fundamentals of Biochemistry (Second Edition), John Wiley and Sons, Inc., New York, NY.
23. Buzayan, J. M., Gerlach, W. L., and Bruening, G. (1986) Non-enzymatic cleavage and ligation of RNAs complementary to a plant virus satellite RNA, *Nature 323*, 349-353.
24. Hutchins, C. J., Rathjen, P. D., Forster, A. C., and Symons R. H. (1986) Self-cleavage of plus and minus RNA transcripts of avocado sunblotch viroid, *Nucleic Acids Research 14*, 3627-3640.
25. Prody, G. A., Bakos, J. T., Buzayan, J. M., Schneider, I. R., and Bruening, G. (1986) Autolytic Processing of Dimeric Plant Virus Satellite RNA, *Science 231*, 1577-1580.
26. Saville, B. J., and Collins, R. A. (1990) A site-specific self-cleavage reaction performed by a novel RNA in Neurospora mitochondria, *Cell 61*, 685-696.
27. Peebles, C. L., Perlman, P. S., Mecklenburg, K. L., Petrillo, M. L., Tabor, J. H., Jarrell, K. A., and Cheng, H. L. (1986) A self-splicing RNA excises an intron lariat, *Cell 44*, 213-223.
28. van der Veen, R., Amberg, A. C., van der Horst, G., Bonen, L., Tabak, H. F., and Grivell, L. A. (1986) Excised group II introns in yeast mitochondria are lariats and can be formed by self-splicing in vitro, *Cell 44*, 225-234.
29. Schmelzer, C., and Schweyen, R. J. (1986) Self-splicing of group II introns in vitro: mapping of the branch point and mutational inhibition of lariat formation, *Cell 46*, 557-565.
30. Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A., and Breaker, R. R. (2004) Control of gene expression by a natural metabolite-responsive ribozyme, *Nature 428*, 281-286.
31. Haugen, P., Simon, D. M., and Bhattacharya, D. (2005) The natural history of group I introns, *Trends in Genetics 21*, 111-119.

32. Rodriguez-Trelles, F., Tarrio, R., and Ayala, F. J. (2006) Origins and evolutions of spliceosomal introns, *Annual Review of Genetics 40*, 47-76.
33. Weeks, K. M., and Cech, T. R. (1995) Protein facilitation of group I intron splicing by assembly of the catalytic core and the 5′ splice site domain, *Cell 82*, 221-230.
34. Valadkhan, S. (2007) The spliceosome: caught in a web of shifting interactions, *Current Opinion in Structural Biology 17*, 310-315.
35. Cannone, J. J., Subramanian, S., Schnare, M. N., Collett, J. R., D'Souza, L. M., Du, Y., Feng, B., Lin, N., Madabusi, L.V., Muller, K. M., Pande, N., Shang, Z., Yu, N., and Gutell, R. R. (2002), The Comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs, *BMC Bioinformatics 3*, 2-32.
36. Michel, F., Jacquier, A., and Dujon, B. (1982) Comparison of fungal mitochondrial introns reveals extensive homologies in RNA secondary structure, *Biochimie 64*, 867-881.
37. Michel, F., and Westhof, E. (1990) Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis, *Journal of Molecular Biology 216*, 585-610.
38. Jaeger, L., Westhof, E., and Michel, F. (1991) Function of P11, a tertiary base pairing in self-splicing introns of subgroup IA, *Journal of Molecular Biology 221*, 1153-1164.
39. Lehnert, V., Jaeger, L., Michel, F., and Westhof, E. (1996) New loop-loop tertiary interactions in self-splicing introns of subgroup IC and ID: a complete 3D model of the Tetrahymena thermophila ribozyme, *Chemistry & Biology 3*, 993-1009.
40. Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A., and Scazzocchio, C. (1982) Making ends meet: a model for RNA splicing in fungal mitochondria, *Nature 300*, 719-724.
41. Waring, R. B., Towner, P., Minter, S. J., and Davies, R. W. (1986) Splice-site selection by a self-splicing RNA of *Tetrahymena*, *Nature 321*, 133-139.
42. Garriga, G., Lambowitz, A. M., Inoue, T., and Cech, T. R. (1986) Mechanism of recognition of the 5' splice site in self-splicing group I introns, *Nature 322*, 86-89.
43. Michel, F., Netter, P., Xu, M. Q., and Shub, D. A. (1990) Mechanism of 3' splice site selection by catalytic core of the *sun*Y intron of bacteriophage T4: the role of a novel base-pairing interaction in group I introns, *Genes and Development 4*, 777-788.
44. Burke, J. M., Esherick, J. S., Burfeind, W. R., and King, J. L. (1990) A 3' splice site binding sequence in the catalytic core of a group I intron, *Nature 344*, 80-82.
45. Tanner, N. K., and Cech, T. R. (1987) Guanosine binding required for cyclization of the self-splicing intervening sequence ribonucleic acid from *Tetrahymena thermophila*, *Biochemistry 26*, 3330-3340.
46. Michel, F., Hanna, M., Green, R., Bartel, D. P., and Szostak, J. W. (1989) The guanosine binding site of the *Tetrahymena* ribozyme, *Nature 342*, 391-395.

47. Been, M. D., and Perrotta, A. T. (1991) Group I intron self-splicing with adenosine: evidence for a single nucleoside-binding site, *Science 252*, 434-437.
48. Barfod, E. T., and Cech, T. R. (1989) The conserved U·G pair in the 5' splice site duplex of a group I intron is required in the first but not the second step of self-splicing, *Molecular and Cellular Biology 9*, 3657-3666.
49. Doudna, J. A., Cormack, B. P., and Szostak, J. W. (1989) RNA structure, not sequence, determines the 5' splice-site specificity of a group I intron, *Proceedings of the National Academy of Sciences, USA 86*, 7402-7406.
50. Pyle, A. M., Moran, S., Strobel, S. A., Chapman, T., Turner, D. H., and Cech, T. R. (1994) Replacement of the conserved G·U with a G-C pair at the cleavage site of the *Tetrahymena* ribozyme decreases binding, reactivity, and fidelity. *Biochemistry 33*, 13856-13863.
51. Knitt, D. S., Narlikar, G. J., and Herschlag, D. (1994) Dissection of the role of the conserved G·U pair in group I RNA self-splicing, *Biochemistry 33*, 13864-13879.
52. Allain, F. H. T., and Varani, G. (1995) Divalent metal ion binding to a conserved wobble pair defining the upstream site of cleavage of group I self-splicing introns, *Nucleic Acids Research 23*, 341-350.
53. Hur, M., and Waring, R. B. (1995) Two group I introns with a C·G basepair at the 5' splice-site instead of the very highly conserved U·G basepair: Is selection post-translational, *Nucleic Acids Research 23*, 4466-4470.
54. Cech, T. R., and Bass, B. L. (1986) Biological catalysis by RNA, *Annual Review of Biochemistry 55*, 599-629.
55. Grosshans, C. A., and Cech, T. R. (1989) Metal ion requirements for sequence-specific endoribonuclease activity of the *Tetrahymena* ribozyme, *Biochemistry 28*, 6888-6894.
56. Pyle, A. M. (1993) Ribozymes: A distinct class of metalloenzymes, *Science 261*, 709-714.
57. Bass, B. L., and Cech, T. R. (1984) Specific interaction between the self-splicing RNA of *Tetrahymena* and its guanosine substrate: implications for biological catalysis by RNA, *Nature 308*, 820-826.
58. Zaug, A. J., and Cech, T. R. (1982) The intervening sequence excised from the ribosomal RNA precursor of *Tetrahymena* contains a 5'-terminal guanosine residue not encoded by the DNA, *Nucleic Acids Research 10*, 2823-2838.
59. Golden, B. L., and Cech, T. R. (1996) Conformational switches involved in orchestrating the successive steps of group I RNA splicing, *Biochemistry 35*, 3754-3763.
60. Price, J. V., and Cech, T. R. (1988) Determinants the 3' splice site for self-splicing of the Tetrahymena pre-rRNA, *Genes and Development 2*, 1439-47.
61. Burke, J. M. (1989) Selection of the 3' splice site in group I introns, *FEBS Letters 250*, 129-33.
62. Suh, E. R., and Waring, R. B. (1990) Base pairing between the 3' exon and an internal guide sequence increases 3' splice site specificity in the *Tetrahymena* self-splicing rRNA intron, *Molecular and Cellular Biology 10*, 2960-2965.

63. Moran, S., Kierzek, R., and Turner, D. H. (1993) Binding of guanosine and 3' splice site analogues to a group I ribozyme: Interactions with functional groups of guanosine and with additional nucleotides, *Biochemistry 32*, 5247-5256.
64. Bevilacqua, P. C., Sugimoto, N., and Turner, D. H., (1996) A mechanistic framework for the second step of splicing catalyzed by the *Tetrahymena* ribozyme, *Biochemistry 35*, 648-658.
65. Russel, R., and Herschlag, D. (1999) Specificity from steric restrictions in the guanosine binding pocket of a group I ribozyme, *RNA 5*, 158-166.
66. Karbstein, K., and Herschlag, D. (2003) Extraordinarily slow binding of guanosine to the *Tetrahymena* group I ribozyme: Implications for RNA reorganization and function, *Proceedings of the National Academy of Science, USA 100*, 2300-2305.
67. Grabowski, P. J., Zaug, A. J., and Cech, T. R. (1981) The intervening sequence of the ribosomal RNA precursor is converted to a circular RNA in isolated nuclei of Tetrahymena, *Cell 23*, 467-476.
68. Zaug, A. J., Grabowski, P. J., and Cech, T. R. (1983) Autocatalytic cyclization of an excised intervening sequence RNA is a cleavage-ligation reaction, *Nature 301*, 578-583.
69. Zaug, A. J., Kent, J. R., and Cech, T. R. (1984) A labile phosphodiester bond at the ligation junction in a circular intervening sequence RNA, *Science 224*, 574-78.
70. Been, M. D., and Cech, T. R. (1987) Selection of circularization sites in a group I IVS RNA requires multiple alignments of an internal template-like sequence, *Cell 50*, 951-961.
71. Zaug, A. J., Grosshans, C. A., Cech, T. R. (1988) Sequence-specific endoribonuclease activity of the *Tetrahymena* ribozyme: Enhanced cleavage of certain oligonucleotide substrates that form mismatched ribozyme-substrate complexes, *Biochemistry 27*, 8924-8931.
72. Zaug, A. J., Davila-Aponte, J A., and Cech, T. R. (1994) Catalysis of RNA cleavage by a ribozyme derived from the group I intron of *Anabaena* pre-$tRNA^{Leu}$, *Biochemistry 33*, 14935-14947.
73. Kuo, L. Y., Davidson, L. A., and Pico, S. (1999) Characterization of the *Azoarcus* ribozyme: tight binding to guanosine and substrate by an unusually small group I ribozyme, *Biochimica et Biophysica Acta 1489*, 281-292.
74. Disney, M. D., Haidaris, C. G., Turner, D. H. (2001) Recognition elements for 5' exon substrate binding to the *Candida albicans* group I intron, *Biochemistry 40*, 6507-6519.
75. Testa, S. M., Haidaris, C. G., Gigliotti, F., and Turner, D. H. (1997) A *Pneumocystis carinii* Group I Ribozyme that Does Not Require 2' OH Groups on its 5' Exon Mimic for Binding to the Catalytic Core, *Biochemistry 36*, 15303-15314.
76. Chowrira, B. M., Berzel-Herranz, A., and Burke, J. M. (1993) Novel RNA polymerization reaction catalyzed by a group I intron-derived ribozyme, *EMBO J. 12*, 3599-3605.

77. Berzal-Herranz, A., Chowrira, B. M., Polsenberg, J. F., and Burke, J. M. (1993) 2' Hydroxyl groups important for exon polymerization and reverse exon ligation reactions catalyzed by a group I ribozyme, *Biochemistry 32*, 8981-8986.
78. Riley, C. A., and Lehman, N. (2003) Generalized RNA-directed recombination of RNA, *Chemistry & Biology 10*, 1233-1243.
79. Woodson, S. A., and Cech, T. R. (1989) Reverse self-splicing of the Tetrahymena group I intron: Implication for the directionality of splicing and for intron transposition, *Cell 57*, 335-45.
80. Roman, J., and Woodson, S. A. (1998) Integration of the Tetrahymena group I intron into bacterial rRNA by reverse splicing in vivo, *Proceedings of the National Academy of Sciences, USA 95*, 2134-39.
81. Sullenger, B. A., and Cech, T. R. (1994) Ribozyme-mediated repair of defective mRNA by targeted, trans-splicing, *Nature 371*, 619-622.
82. Jones, J. T., Lee, S. W., and Sullenger, B. A. (1996) Tagging ribozyme reaction sites to follow trans-splicing in mammalian cells, *Nature Medicine 2*, 643-648.
83. Kohler, U., Ayre, B. G., Goodman, H. M., and Haseloff, J. (1999) Trans-splicing ribozymes for targeted gene delivery, *Journal of Molecular Biology 285*, 1935-1950.
84. Watanabe, T., and Sullenger, B. A. (2000) Induction of wild-type p53 activity in human cancer cells by ribozymes that repair mutant p53 transcripts, *Proceedings of the National Academy of Sciences, USA 97*, 8490-8494.
85. Lan, N., Rooney, B. L., Lee, S. W., Howrey, R. P., Smith, C. A., and Sullenger, B. A. (2000) Enhancing RNA repair efficiency by combining trans-splicing ribozymes that recognize different accessible sites on a target RNA, *Molecular Therapy 2*, 245-255.
86. Ayre, B. G., Kohler, U., Turgeon, R., and Haseloff, J. (2002) Optimization of trans-splicing ribozyme efficiency and specificity by in vivo genetic selection, *Nucleic Acids Research 30*, e141.
87. Byun, J., Lan, N., Long, M., and Sullenger, B. A. (2003) Efficient and specific repair of sickle beta-globin RNA by trans-splicing ribozymes, *RNA 9*, 1254-1263.
88. Shin, K. S., Sullenger, B. A., and Lee, S. W. (2004) Ribozyme-mediated induction of apoptosis in human cancer cells by targeted repair of mutant p53 RNA, *Molecular Therapy 10*, 365-372.
89. Lundblad, E. W., Haugen, P., and Johansen, S. D. (2004) Trans-splicing of a mutated glycosylasparaginase mRNA sequence by a group I ribozyme, *European Journal of Biochemistry 271*, 4932-4938.
90. Fiskaa, T., Lundblad, E. W., Henriksen, J. R., Johansen, S. D., and Einvik, C. (2006) RNA reprogramming of mannosidase mRNA sequences in vitro by myxomycete group IC1 and IE ribozymes, *The FEBS Journal 273*, 2789-2800.
91. Alexander, R. C., Baum, D. B., and Testa, S. M. (2005) 5' Transcript replacement in vitro catalyzed by a group I intron-derived ribozyme, *Biochemistry 44*, 7796-7804.
92. Milligan, J. F., Groebe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates, *Nucleic Acids Research 15*, 8783-98.

93. Tsai, C. H., and Dreher, T. W. (1993) *In vitro* transcription of RNAs with defined 3' termini from PCR-generated templates, *Biotechniques 14*, 58-61.
94. Suh, E. and Waring, R. B. (1993) The conserved terminal guanosine of a group I intron can help prevent reopening of the ligated exons, *Journal of Molecular Biology 232*, 375-85.
95. Studier, F. W., Rosenberg, A. H., and Dunn, J. J. (1990) Use of T7 RNA polymerase to direct expression of cloned genes, *Methods in Enzymology 185*, 60-89.
96. Fuerst, T. R., and Moss, B. (1989) Structure and stability of mRNA synthesized by vaccinia virus-encoded bacteriophage T7 RNA polymerase in mammalian cells, *Journal of Molecular Biology 206*, 333-48.
97. Zaug, A. J., Kent, J. R. and Cech, T. R. (1985) Reactions of the intervening sequence of the *Tetrahymena* ribosomal ribonucleic acid precursor pH dependence of cyclization and site-specific hydrolysis. *Biochemistry*, *24*, 6211-6218.
98. Johnson, A. J., Baum, D. B., Tye, J., Bell, M. A., and Testa, S. M. (2003) Molecular recognition properties of IGS-mediated reactions catalyzed by a *Pneumocystis carinii* group I intron, *Nucleic Acids Research 31*, 1921-1934.
99. Grosshans, C. A., and Cech, T. R. (1991) A hammerhead ribozyme allows synthesis of a new form of the Tetrahymena ribozyme homogeneous in length with a 3' end blocked for transesterification, *Nucleic Acids Research 19*, 3875-80.
100. Ritchings, B. W., and Lewin, A. S. (1992) Mutational evidence for competition between the P1 and the P10 helices of a mitochondrial group I intron, *Nucleic Acids Research 20*, 2349-53.
101. Nielsen, H., Fiskaa, T., Birgisdottir, A. B., Haugen, P., Einvik, C., and Johansen, S. (2003) The ability to form full-length intron RNA circles is a general property of nuclear group I introns, *RNA 9*, 1464-1475.
102. Herschlag, D, and Cech, T. R. (1990) Catalysis of RNA cleavage by the Tetrahymena thermophila ribozyme. 2. Kinetic description of the reaction of an RNA substrate that forms a mismatch at the active site, Biochemistry 29, 10172-10180.
103. Pyle, A. M., McSwiggen, J. A., and Cech, T. R. (1990) Direct measurement of oligonucleotide substrate binding to wild-type and mutant ribozymes from Tetrahymena, *Proceedings of the National Academy of Sciences, USA 87*, 8187-8191.
104. Fedor, M. J., and Uhlenbeck, O. C. (1992) Kinetics of intermolecular cleavage by hammerhead ribozymes, *Biochemistry 31*, 12042-12054.
105. Franzen, J. S., Zhang, M., and Peebles, C. L. (1993) Kinetic analysis of the 5' splice junction hydrolysis of a group II intron promoted domain 5, *Nucleic Acids Research 21*, 627-634.
106. Smith, D., and Pace, N. R. (1993) Multiple magnesium ions in the ribonuclease P reaction mechanism, *Biochemistry 32*, 5273-5281.
107. Pyle, A. M., and Green, J. B. (1994) Building a kinetic framework for group II intron ribozyme activity: Quantitation of interdomain binding and reaction rate, *Biochemistry 33*, 2716-2725.

108. Hertel, K. J., Herschlag, D., and Uhlenbeck, O. C. (1994) A kinetic and thermodynamic framework for the hammerhead ribozyme reaction, *Biochemistry 33*, 3374-3385.
109. Esteban, J. A., Banerjee, A. R., and Burke, J. M. (1997) Kinetic mechanism of the hairpin ribozyme. Identification and characterization of two nonexchangeable conformations, *Journal of Biological Chemistry 272*, 13692-13639.
110. Bergman, N. H., Johnston, W. K., and Bartel, D. P. (2000) Kinetic framework for ligation by an efficient RNA ligase ribozyme, *Biochemistry 39*, 3115-3123.
111. Shih, I., and Been, M. D. (2000) Kinetic scheme for intermolecular RNA cleavage by a ribozyme derived from Hepatitis delta virus RNA, *Biochemistry 39*, 9055-9066.
112. Rose, I. A., O'Connell, E. L., and Litwin, S. (1974) Determination of the rate of hexokinase-glucose dissociation by the isotope-trapping method, *Journal of Biological Chemistry 249*, 5163-5168.
113. Weeks, K. M., and Crothers, D. M. (1992) RNA binding assays for Tat-derived peptides: Implications for specificity, *Biochemistry 31*, 10281-10287.
114. McConnell, T. S., Cech, T. R., and Herschlag, D. (1993) Guanosine binding to the Tetrahymena ribozyme: Thermodynamic coupling with oligonucleotide binding, *Proceedings of the National Academy of Sciences, USA 90*, 8362-8366.
115. Dahm, S. C., Derrick, W. B., and Uhlenbeck, O. C. (1993) Evidence for the role of solvated metal hydroxide in the hammerhead cleavage mechanism, *Biochemistry 32*, 13040-13045.
116. Curtis, E. A., and Bartel, D. P. (2001) The hammerhead cleavage reaction in monovalent cations, *RNA 7*, 546-552.
117. Dahm, S. C., and Uhlenbeck, O. C. (1991) Role of divalent metal ions in the hammerhead RNA cleavage reaction, *Biochemistry 30*, 9464-9469.
118. Viola, R. E., and Cleland, W. W. (1978) Use of pH studies to elucidate the chemical mechanism of yeast hexokinase, *Biochemistry 17*, 4111-4117.
119. Knitt, D. S. and Herschlag, D. (1996) pH dependencies of the Tetrahymena ribozyme reveal an unconventional origin of an apparent pKa, *Biochemistry 33*, 1560-1570.
120. Herschlag, D., and Kosla, M. (1994) Comparison of pH dependencies of the Tetrahymena ribozyme reactions with RNA 2'-substituted and phosphorothioate substrates reveal a rate-limiting conformational step, *Biochemistry 33*, 5291-5297.
121. Saenger, W. (1988) *Principles of nucleic acid structure*, 2nd edn. Springer-Verlag GmbH & Co KG, Berlin, Germany.
122. Fersht, A. (1999) *Structure and mechanism in protein science: A guide to enzyme catalysis and protein folding*, W. H. Freeman, New York.
123. Eigen, M., and Hammes, G. G. (1963) Elementary steps in enzyme reactions (As studied by relaxation spectroscopy), *Adv. Enzymol. Relat. Areas Mol. Biol. 25*, 1-38.
124. Quast, U., Engel, J., Heumann, H., Krause, G., and Steffen, E. (1974) Kinetics of the interaction of bovine pancreatic trypsin inhibitor (Kunitz) with alpha-chymotrypsin, *Biochemistry 13*, 2512-2520.
125. Hammes, G. G., and Hurst, J. K. (1969) Relaxation spectra of adenosine triphosphate-creatine phosphotransferase, *Biochemistry 8*, 1083-1094.

126. Czerlinski, G. H., and Schreck, G. (1964) Chemical relaxation of the reaction of malate dehydrogenase with reduced nicotinamide adenine dinucleotide determined by fluorescence detection, *Biochemistry 3*, 89-100.
127. Porschke D & Eigen M (1971) Co-operative non-enzymatic base recognition. 3. Kinetics of the helix-coil transition of the oligoribouridylic—oligoriboadenylic acid system and of oligoriboadenylic acid alone at acidic pH, *Journal of Molecular Biology 62*, 361-381.
128. Nelson, J. W., and Tinoco, I. Jr (1982) Comparison of the kinetics of ribooligonucleotide, deoxyribooligonucleotide, and hybrid oligonucleotide double-strand formation by temperature-jump kinetics, *Biochemistry 21*, 5289-5295.
129. Craig, M. E., Crothers, D. M., and Doty, P. (1971) Relaxation kinetics of dimer formation by self complementary oligonucleotides, *Journal of Molecular Biology 62*, 383-401.
130. Breslauer, K. J., and Bina-Stein, M. (1977) Relaxation kinetics of the helix-coil transition of a self-complementary ribo-oligonucleotide: A7U7, *Biophysical Chemistry 7*, 211-216.
131. Ravetch, J., Gralla, J., and Crothers, D. M. (1974) Thermodynamic and kinetic properties of short RNA helices: The oligomer sequence AnGCUn, *Nucleic Acids Research 1*, 109-127.
132. Bevilacqua, P. C., Kierzek, R., Johnson, K. A., and Turner, D. H. (1992) Dynamics of ribozyme binding of substrate revealed by fluorescence-detected stopped-flow methods, *Science 258*, 1355-1358.
133. Santoro, S. W., and Joyce, G. F. (1997) A general purpose RNA-cleaving DNA enzyme, *Proceedings of the National Academy of Sciences, USA 94*, 4262-4266.
134. Li, Y., and Breaker, R. R. (1999) Kinetics of RNA degradation by specific base catalysis of transesterification involving the 2'-hydroxyl group, *Journal of the American Chemical Society 121*, 5364-5372.
135. Narlikar, G. J., Gopolakrishnan, V., McConnell, T. S., Usman, N. and Herschlag, D. (1995) Use of binding energy by an RNA enzyme for catalysis by positioning and substrate destabilization, *Proceedings of the National Academy of Sciences, USA 92*, 3668-3672.
136. Bass, B. L. and Cech, T. R. (1986) Ribozyme inhibitors: Deoxyguanosine and dideoxyguanosine are competitive inhibitors of self-splicing of the Tetrahymena ribosomal ribonucleic acid precursor, *Biochemistry 25*, 4473-4477.
137. Shan, S., Kravchuk, A. V., Piccirilli, J. A., and Herschlag, D. (2001) Defining the catalytic metal ion interactions in the Tetrahymena ribozyme reaction, *Biochemistry 40*, 5161-5171.
138. Shan, S., Yoshida, A., Sun, S., Piccirilli, J. A., and Herschlag, D. (1999) Three metal ions at the active site of the Tetrahymena group I ribozyme, *Proceedings of the National Academy of Sciences, USA 96*, 12299-12304.
139. Sjogren, A. S., Pettersson, E., Sjoberg, B. M., and Stromberg, R. (1997) Metal ion interaction with cosubstrate in self-splicing of group I introns, *Nucleic Acids Research 25*, 648-653.

140. Stahley, M. R., and Strobel, S. A. (2005) Structural evidence for a two-metal-ion mechanism of group I intron splicing, *Science 309*, 1587-1590.
141. Adams, P. L., Stahley, M. R., Gill, M. L., Kosek, A. B., Wang, J., and Strobel, S. A. (2004) Crystal structure of a group I intron splicing intermediate, *RNA 10*, 1867-1887.
142. Karbstein, K., Lee, J. and Herschlag, D. (2007) Probing the role of a secondary structure element at the 5'- and 3-splice sites in group I intron self-splicing: The *Tetrahymena* L-16 ScaI ribozyme reveals a new role of the G•U pair in self-splicing. *Biochemistry*, *46*, 4861-4875.
143. Sargueil, B., and Tanner, N. K. (1993) A shortened form of the *Tetrahymena thermophila* group I intron can catalyze the complete self-splicing reaction *in trans*, *Journal of Molecular Biology 233*, 629-643.
144. Drude, I., Vauleon, S., and Muller, S. (2007) Twin ribozyme mediated removal of nucleotides from an internal RNA site, *Biochemical and Biophysical Research Communications 363*, 24-29.
145. Grasby, J. A., and Gait, M. J. (1994) Synthetic oligoribonucleotides carrying site-specific modifications for RNA structure-function analysis, *Biochimie 76*, 1223-1234.
146. Verma, S., and Eckstein, F. (1998) Modified oligonucleotides: Synthesis and strategy for users, *Annual Review of Biochemistry 67*, 99-134.
147. Das, S. R., Fong, R., and Piccirilli, J. A. (2005) Nucleotide analogues to investigate RNA structure and function, *Current Opinion in Chemical Biology 9*, 1-9.
148. Shabarova, Z. A. (1988) Chemical development in the design of oligonucleotide probes for binding to DNA and RNA, *Biochimie 70*, 1323-1334.
149. Moore, M. J. and Sharp, P. A. (1992) Site-specific modification of pre-mRNA: The 2' hydroxyl groups at the splice sites, *Science*. *256*, 992-997.
150. Bain, J. D. and Switzer, C. (1992) Regioselective ligation of oligoribonucleotides using DNA splints, *Nucleic Acids Research 20*, 4372.
151. Milligan, J. F., and Uhlenbeck, O C. (1989) Synthesis of small RNAs using T7 RNA polymerase, *Methods in Enzymology 180*, 51-62.
152. Aurup, H., Williams, D. M., and Eckstein, F. (1992) 2'-fluoro- and 2'-amino-2'-deoxynucleoside 5'-triphosphates as substrates for T7 RNA polymerase, *Biochemistry 31*, 9636-9641.
153. Hirao, I., Ohtsuki, T., Fujiwara, T., Mitsui, T., Yokogawa, T., Okuni, T., Nakayama, H., Takio, K., Yabuki, T., Kigawa, T., Kodama, K., Yokogawa, T., Nishikawa, K., and Yokoyama, S. (2002) An unnatural base pair for incorporating amino acid analogs into proteins, *Nature Biotechnology 20*, 177-182.
154. Mitsui, T., Kimoto, M., Harada, Y., Yokoyama, S., and Hirao, I. (2005) An efficient unnatural base pair for a base-pair expanded transcription system, *Journal of the American Chemical Society 127*, 8652-8658.
155. Kawai, R., Kimoto, M., Ikeda, S., Mitsui, T., Endo, M., Yokoyama, S. and Hirao, I. (2005) Site-specific fluorescent labeling of RNA molecules by specific transcription using unnatural base pairs, *Journal of the American Chemical Society 127*, 17286-17295.

156. Vauleon, S., Ivanov, S. A., Gwiazda, S., and Muller, S. (2005) Site-specific fluorescent and affinity labeling of RNA by using a small engineered twin ribozyme, *ChemBioChem 6*, 2158-2162.
157. Baum, D. A., and Silverman, S. K. (2007) Deoxyribozyme-catalyzed labeling of RNA, *Angewandte Chemie International Edition (English) 46*, 3502-3504.
158. Romaniuk, P. J., and Uhlenbeck, O. C. (1983) Joining of RNA molecules with RNA ligase, *Methods in Enzymology 100*, 52-59.
159. Maunders, M. J. (1993) DNA and RNA ligases, *Methods in Molecular Biology 16*, 213-230.
160. Moore, M. J., and Query, C. C. (2000) Joining of RNAs by splinted ligation, *Methods in Enzymology 317*, 109-123.
161. Stark, M. R., Pleiss, J. A., Deras, M., Scaringe, S. A., and Rader, S. D. (2006) An RNA ligase-mediated method for the efficient creation of large, synthetic RNAs, *RNA 12*, 2014-2019.

Vita

Perry Patrick Dotson II was born in Charleston, West Virginia on May 27, 1974. In 1997, he obtained his bachelor's degree in biology at Marshall University. He also obtained a bachelor's degree in chemistry in 2000 at West Virginia State College (now West Virginia State University). After graduation, he was employed first as a laboratory technician for Union Carbide Corporation and then as a chemical technologist for The Dow Chemical Company. He left after two years to pursue graduate work in chemistry at the University of Kentucky in the laboratory of Dr. Stephen M. Testa.

List of Publications:

1. Dotson II, P. P., and Testa, S. M. (2006) Group I intron-derived ribozyme recombination reactions, *Recent Developments in Nucleic Acids Research 2*, 307-324.

2. Dotson II, P. P., Sinha, J., and Testa, S. M. (2008) A *Pneumocystis carinii* group I intron-derived ribozyme utilizes an endogenous guanosine as the first reaction step nucleophile in the trans excision-splicing reaction, *Biochemistry 47*, 4780-4787.

3. Dotson II, P. P.*, Sinha, J.*, and Testa, S. M. (2008) Kinetic characterization of the first step of ribozyme-catalyzed trans excision-splicing reaction, *The FEBS Journal*, in press.

4. Dotson II, P. P., Frommeyer, K. N., and Testa, S. M. (2008) Sequence-specific insertion of chemically modified oligonucleotides into RNA substrates via the ribozyme-mediated trans insertion-splicing reaction, *ChemBioChem*, submitted for publication.

5. Dotson II, P. P., Johnson, A. K., and Testa, S. M. (2008) Group I intron-derived ribozymes from *Tetrahymena thermophila* and *Candida albicans* catalyze the trans excision-splicing reaction, *Nucleic Acids Res*., manuscript in preparation.